



UNIVERSIDAD DE MÁLAGA

ETSI DE TELECOMUNICACIÓN

Correlaciones en series módulo y signo. Aplicación a medidas de no-linealidad en series cardíacas

Manuel Gómez Extremera

Tesis doctoral dirigida por: Pedro A. Bernaola Galván

Programa: Ingeniería Mecánica y Eficiencia Energética


Departamento de Física Aplicada II

Málaga, diciembre de 2017



UNIVERSIDAD
DE MÁLAGA

AUTOR: Manuel Gómez Extremera

 <http://orcid.org/0000-0003-4982-3481>

EDITA: Publicaciones y Divulgación Científica. Universidad de Málaga



Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional:

<http://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>

Cualquier parte de esta obra se puede reproducir sin autorización
pero con el reconocimiento y atribución de los autores.

No se puede hacer uso comercial de la obra y no se puede alterar, transformar o hacer
obras derivadas.

Esta Tesis Doctoral está depositada en el Repositorio Institucional de la Universidad de
Málaga (RIUMA): riuma.uma.es

Agradecimientos

En primer lugar a mi director de tesis, el Dr. Pedro A. Bernaola Galván. Gracias por haber confiado en mí para tu proyecto, por todo lo que he aprendido de ti día tras día, por haberme ayudado a crecer en todos los ámbitos y por haber tenido siempre que lo he necesitado tiempo para mí. Al Dr. Pedro J. Carpena Sánchez, por su ayuda desinteresada a lo largo de estos años y por haber contribuido también de forma fehaciente día tras día a mi formación académica y personal. Al resto de miembros del grupo de Física Estadística de Sistemas Complejos y Biocomputación (Dra. Ana V. Coronado y ‘Conchita’) por haber estado ahí siempre que os he necesitado. También agradecer al resto de profesores y compañeros del Departamento de Física Aplicada II con los que he tenido el placer de compartir docencia y a Laureano, por toda su ayuda en estos últimos meses.

A los miembros del grupo de redes fisiológicas de la Universidad de Boston, en especial al Prof. Plamen Ivanov, por haberme dado la oportunidad de trabajar junto a él en un centro de investigación de primer nivel mundial. A Luis por nuestras charlas no tan científicas, Fabrizio, Luca, Wanting, Xiyun, etc.

A mis compañeros de licenciatura de Granada, en especial a José (Dr.) por la ayuda prestada innumerables veces, a Carmen, Pilar, Luis y ‘Karry’.

A mis compañeros del Dpto de Física Atómica, Molecular y Nuclear de la Universidad de Granada, ‘Wil’, Juan, Katerine, Jordi, Alex, Rodrigo, ‘Moi’, Javi y a los Doctores Ángel V. Delgado, Antonio M. Lallena y Marta Anguiano por haberme dado la oportunidad de iniciarme en el mundo de la investigación.

Por último, no me gustaría olvidarme de aquellas personas que fuera del ámbito académico, también han contribuido de manera notable. En especial a mi grupo de amigos de toda la vida, y por supuesto a toda mi familia, sobre todo mis padres y mi hermana, por haberme dado todo y apoyarme siempre; mi tío Laureano, por inculcarme interés por la ciencia y mi abuela, aunque nunca pueda explicarte en qué consiste mi investigación.

¡Gracias a todos!

Índice general

Lista de abreviaturas	1
Summary	3
Introducción	21
1. Métodos para el estudio y análisis de series temporales	25
1.1. Serie Temporal	26
1.1.1. Medidas de dependencia: covarianza y autocorrelación	26
1.1.2. Estacionariedad	27
1.2. Ruido Fraccionario Gaussiano y Movimiento Fraccionario Browniano	28
1.3. Técnicas de Análisis de Fluctuación	30
1.3.1. Fluctuation Analysis	31
1.3.2. Detrended Fluctuation Analysis	31
1.4. Método de Filtrado de Fourier	34
1.4.1. Generación de ruidos fraccionarios con espectro de potencias limpio	35
1.5. Multifractal DFA	37
1.6. Transformada Wavelet	39
1.7. Randomización de fases	41
1.8. Transformación a distribución Gaussiana	41
2. Módulo y signo de series con correlaciones de largo alcance	45
2.1. Descomposición	47
2.2. Composición mediante series módulo y signo	53
2.2.1. Caso $\alpha_{\text{sign}} < 1$	54
2.2.2. Caso $\alpha_{\text{mag}} < 1, \alpha_{\text{sign}} > 1$	56
2.2.3. Caso $\alpha_{\text{sign}} > 1, \alpha_{\text{mag}} > 1$	58
2.3. Propiedades multifractales de series compuestas	62

3. Función de autocorrelación de series módulo y signo de fGns	65
3.1. Autocorrelación de la serie módulo para ruidos fraccionarios Gaussianos . .	66
3.2. Autocorrelación de la serie cuadrática	74
3.3. Autocorrelación de la serie signo	80
3.4. Modelo no lineal de series	88
3.4.1. No-linealidad y multifractalidad	93
3.4.2. Efectos de tamaño finito en medidas de no-linealidad	95
4. Medidas de no-linealidad en series cardíacas	99
4.1. Estudio de no-linealidad durante ejercicio y reposo	100
4.2. Estudio de no-linealidad en sujetos con entrenamiento aeróbico y de fuerza	110
4.3. Comparación de resultados mediante el exponente α_1 del DFA	111
4.4. Otros resultados preliminares	116
5. Conclusions	119
5.1. Chapter 1	119
5.2. Chapter 2	119
5.3. Chapter 3	121
5.4. Chapter 4	122
5.5. General conclusion and future perspective	124
A. Fluctuaciones del signo en series no estacionarias ($\alpha \geq 1$)	125
B. Función de autocorrelación de series temporales con módulo y signo desacoplados.	129
C. Conceptos de Fisiología	131
C.1. Sistema Nervioso Autónomo	131
C.2. Latido Cardíaco	132
C.2.1. Observables cardíacos	134
D. Publicaciones y contribuciones	137
D.1. Publicaciones	137
D.2. Contribuciones	138
Bibliografía	139

Índice de figuras

1.1. Ejemplo de FA y espectro de potencias para un fGn correlacionado	32
1.2. Ejemplo de FA y espectro de potencias para un fGn anticorrelacionado . .	32
1.3. Ejemplo de dos series temporales con igual exponente de correlación α y distinto scaling	34
1.4. Ejemplos de ruidos Gaussianos generados mediante el método de filtrado de Fourier	36
1.5. Espectros de potencias de series sintéticas de tamaño 2^{14} generadas me- diante el método expuesto en (de Moura and Lyra, 1998)	38
1.6. Ejemplo de espectros multifractales obtenidos mediante el MFDFA	39
1.7. Ejemplo de serie sintética y transformada Wavelet mediante primera deri- vada Gaussiana	41
1.8. Randomización de fases en series con y sin propiedades no lineales	42
1.9. Transformación de series con distribución no Gaussiana a distribución Gaus- siana	43
1.10. DFAs de una serie antes y después transformar su distribución de valores a Gaussiana	43
2.1. Ejemplo de serie correlacionada obtenida mediante el FFM	47
2.2. Efectos de tamaño finito en ruidos Gaussianos obtenidos mediante el FFM	48
2.3. Exponentes de correlación de módulo y signo obtenidos mediante el DFA en función de α_{in}	49
2.4. Anticorrelaciones de corto alcance en series signo	50
2.5. Efecto del acoplamiento para obtención de anticorrelaciones	50
2.6. Ejemplo de serie compuesta generada mediante la multiplicación de módulo y signo obtenidos de series originales independientes para $\alpha_{sign} < 1$	55
2.7. Ejemplos de series compuestas generadas mediante la multiplicación de módulos y signos procedentes de distintas series para el caso $\alpha_{mag} < 1$ y $\alpha_{sign} > 1$	56

2.8. Posición del crossover ℓ_c en función del tamaño de la secuencia N	58
2.9. Ejemplo de series compuestas generadas a través de la multiplicación de módulos y signos procedentes de series independientes para $\alpha_{\text{mag}} > 1$ y $\alpha_{\text{sign}} > 1$	60
2.10. Distribución de exponentes para series obtenidas con $\alpha_{\text{mag}} \simeq 1.9$ y $\alpha_{\text{sign}} \simeq 1.5$	61
2.11. Propiedades multifractales de series artificiales generadas mediante la com- posición de módulos y signos independientes	63
3.1. Autocorrelación de la serie módulo $C_{ x }$ en función de la autocorrelación de la serie original C_x	73
3.2. Función de autocorrelación C_x y función de autocorrelación de la serie módulo $C_{ x }$ para dos fGns (uno con $H > 0.75$ y otro con $H < 0.75$)	75
3.3. Autocorrelación de la serie módulo $C_{ x }$ y la serie cuadrática C_{x^2} en función de la autocorrelación de la serie original C_x	79
3.4. Autocorrelación de la serie signo C_s en función de la autocorrelación de la serie original C_x	87
3.5. Función de autocorrelación en función de la distancia ℓ de serie original C_x y serie signo C_s para dos fGns con $\gamma = 0.3$ y $\gamma = 1.4$	89
3.6. Autocorrelación de la serie módulo $C_{ x }$ en función de la autocorrelación de la serie signo C_s para ruidos lineales	90
3.7. Autocorrelación de la serie módulo $C_{ x }$, en función de la autocorrelación de la serie original C_x para series no lineales	92
3.8. Índice de no-linealidad Δ en función de H_1 para series generadas mediante composición de módulos y signos independientes	95
3.9. Índice de no-linealidad Δ en función del tamaño de la secuencia en ruidos fraccionarios Gaussianos.	97
3.10. Índice de no-linealidad Δ en función del tamaño de la secuencia en series no lineales sintéticas	98
4.1. Ejemplo de serie cardíaca en reposo y ejercicio moderado de un futbolista aficionado	105
4.2. Autocorrelación de señal incremento RR y señal módulo de un futbolista aficionado	106
4.3. Índice de no-linealidad para 10 futbolistas aficionados	107
4.4. Autocorrelación de señal incremento RR y señal módulo de un futbolista profesional	108
4.5. Índice de no-linealidad para 12 futbolistas profesionales	109

4.6. Promedio del índice de no-linealidad para futbolistas profesionales y aficionados durante reposo y ejercicio	110
4.7. Índice de no-linealidad Δ para futbolistas y culturistas	114
4.8. Comparación de datos fisiológicos mediante exponente α_1	115
4.9. Ejemplo de subida y bajada del ritmo cardíaco	116
4.10. Comparación del índice de no-linealidad durante ejercicio y recuperación	118
C.1. Configuración de electrodos en un equipo de medida BIOPAC	133
C.2. Ejemplo de ciclo cardíaco en un sujeto sano	133
C.3. Esquema del corazón humano	134

Índice de Tablas

1. Results obtained for composition by means of independent magnitude and sign time series.	9
2.1. Resultados obtenidos para la composición de series a través de módulos y signos independientes.	61
4.1. Medidas de variabilidad cariaca (HRV) durante el reposo en futbolistas aficionados.	102
4.2. Datos de futbolistas profesionales durante el reposo.	109
4.3. Medidas de variabilidad cardíaca (reposo) en futbolistas (profesionales y aficionados).	112
4.4. Medidas de variabilidad cardíaca (reposo) en culturistas (aficionados). . . .	113

Lista de abreviaturas

FA Análisis de Fluctuación (Fluctuation Analysis)
DFA Análisis de Fluctuación sin tendencia (Detrended Fluctuation Analysis)
MFDFA Análisis multifractal de fluctuación sin tendencia (Multifractal Detrended Fluctuation Analysis)
FFM Método de Filtrado de Fourier (Fourier Filtering Method)
fGn ruido fraccionario Gaussiano (fractional Gaussian noise)
fBm movimiento fraccionario Browniano (fractional Brownian motion)
RR período entre latidos (interbeat interval)
NSA Sistema Nervioso Autónomo (Autonomus Nervious System)
mag módulo (magnitude)
sign signo (sign)

Summary

Introduction

A great variety of phenomena associated with different fields such as Physics, Physiology, Economy, Biology, Geology, Meteorology, etc, show output signals which may appear as erratic and noisy, but in fact, possess long-range correlations and fractal structure (Mandelbrot, 1975). Some examples where long-range correlations have been detected are: human DNA (Hackenberg et al., 2011, 2012; Li and Kaneko, 1992; Peng et al., 1992; Voss, 1992), music (Voss and Clarke, 1978), Hidrology (Hurst, 1951), human heartbeat series (Peng et al., 1993b), keyword detection in texts (Carpena et al., 2009; Ortuno et al., 2002), Economy (Peters, 1994), etc. Moreover, in some cases, the value of correlation is linked to the state of the system, e.g. in the human heartbeat series the correlations change dramatically from healthy to pathological subjects (Ivanov, 1999a), in different physiological states (Ivanov et al., 1999) or as a function of age (Iyengar et al., 1996). However, there also are some systems where series with identical long-range correlations present totally different dynamical properties (Ashkenazy, 2001). These systems present complex interactions (quadratic, exponential, trigonometric, etc) far away from linearity, thus, an extensive study of these interactions can shed light on the dynamics of the system. Nevertheless, it is very difficult to find out the functional dependence of these interactions, and the goal, in most of the cases, is just to assess the degree of complexity or nonlinearity in the system (Ivanov, 1999a; Kalisky et al., 2005). A simple approach to break this degeneration is to analyze separately the correlation properties of magnitude (absolute value of the series) and sign series (+1 if the series is positive and -1 if it is negative) due to: (i) correlations in the magnitude time series have widely been associated to nonlinear and multifractal properties (Ashkenazy, 2001, 2003; Gómez-Extremuera et al., 2016; Kalisky et al., 2005) and (ii) the correlations in the sign series control the linear correlations of the signal (Ashkenazy, 2001, 2003; Carretero-Campos et al., 2012). Thus, studying the correlation properties of magnitude and sign series can be of great utility to better understand the coupling mechanisms between both signals and consequently,

obtaining information about complex systems where the dynamics is controlled by two competitive forces, e.g. in the human heart the increases and decreases or cardiac rhythm are controlled by Sympathetic and Parasympathetic branches of the Autonomic Nervous System (NSA).

The standard method to quantify correlations is DFA (Detrended Fluctuation Analysis) (Peng et al., 1994b). This method removes trends in the time series under study and it is able to estimate (indirectly) the exponent of correlation in power-law¹ correlated time series. However, DFA presents an important drawback: when applying to real data, in most of the cases, fluctuation does not follow a power-law as a function of the scale, so the use of DFA could lead to spurious results (it is not clear yet what deviations over the power-law mean). Furthermore, it requires that the series is long enough in order to get good statistics. For the magnitude and sign series, even in the cases where these requirements are fulfilled, DFA is not able to accurately estimate correlations under certain circumstances (Carpena et al., 2017). In order to solve this problem, an interesting option is the study of the autocorrelation function itself, which despite being noisy and sensitive to the size of the series, provides us with the exact information we are looking for.

Then, the objectives of this work are the study of magnitude and sign properties by means of DFA and autocorrelation function, and then, detecting possible spurious results given by DFA. Moreover, based on the widely known fact that given a time series, correlations in its magnitude series are linked to the nonlinear properties, we will propose a new measure of nonlinearity obtained from the autocorrelation function. Finally, we want to apply this new measure of nonlinearity to real data (heartbeat) in different conditions and physiological states.

Magnitude and sign of long-range correlated time series

As we mentioned in the introduction of this summary, the study of magnitude and sign is an useful technique to obtain information from systems with identical linear correlations, but different nonlinear properties (Ashkenazy, 2001). Moreover, from an intuitive point of view, magnitude and sign series also contain complementary information about the original series (time series from which magnitude and sign are obtained). Magnitude measures how big the changes are, whilst sign measures their direction. As relevant examples of the utility of magnitude and sign it is worth-mentioning the human heart-

¹When a time series is power-law correlated, the fluctuation also follows a power-law dependence as a function of the scale.

beat fluctuation (see Introduction), Fluid Dynamics (Zhu, 2012), Geology (Makse et al., 1996b,c), Geophysics (Bartos and János, 2006; Li et al., 2014) and Economy (Liu, 1999). However, despite how important magnitude and sign are in order to fully understand the behavior of complex systems, there are still open questions: for example, given a time series with known long-range correlations it is important to know if there are also correlations present in the magnitude and sign series (decomposition problem from now on). Moreover, time series present in real data from different systems usually show a totally different behavior, thus, a systematic study of the decomposition problem might be an struggling task. Then, instead of studying real time series, we will study synthetic time series (fractional Gaussian noises and fractional Brownian motions, see Sec. 1.2 for a detailed description of these models) which are commonly used in order to model the behavior of natural series. Another open question is how is the coupling between magnitude and sign to conceive the product signal, i.e. every time we multiply the magnitude series by the sign series we obtain a product signal (in case we decompose a series into magnitude and sign, and then, multiply magnitude by sign, we will obtain again the original signal). Obviously, a systematic study of the coupling between magnitude and sign would be of great interest to improve the understanding of the mechanisms which control both. Here, we are interested in how the magnitude and/or sign control the correlations in the product series, nevertheless, magnitude and sign will have different coupling in time series coming from different systems. Then, we can confront this problem from another point of view: we can systematically study the correlations of time series generated by means of uncoupled magnitude and sign (composition problem from now on). The composition can be useful to understand the behavior of complex systems characterized by the coupling of two different mechanisms, one controlling the dynamics of the magnitude and the other, the sign's.

Decomposition

Here, we quantify linear correlations in magnitude and sign series obtained by means of the decomposition of linear Gaussian long-range correlated fractal time series (obtained using the FFM, see Sec. 1.4). The procedure can be summarized as follows: first, we generate time series with exponent of correlation $\alpha_{in} \in [0, 2]$, for each series we calculate its magnitude and sign series, and measure the linear correlations of the magnitude and sign series (all measures of correlations are by means of DFA). In Fig. 2.3 we show the results. We can mainly distinguish three regions:

- (i) $\alpha_{in} \leq 0.5$. Despite the anticorrelations in the time series, both magnitude and sign are essentially uncorrelated. Taking this into account the anticorrelations present in

the time series are due to the coupling between magnitude and sign. Thus, by means of decomposition is not possible to generate anticorrelated time series (measured by DFA). This limitation has also been found in other methods (Izrailev et al., 2007; Keshet, 2004; Usatenko et al., 2014).

- (ii) $0.5 < \alpha_{\text{in}} < 1$. The sign series present correlations in the whole interval, whilst the magnitude is only correlated for values $\alpha_{\text{in}} > 0.75$. The correlations in the series are controlled by those in the sign, no matter if the magnitude is correlated or not. Moreover, taking into account the analytic relation found in (Apostolov et al., 2008):

$$C(\ell) = \sin \left[\frac{\pi}{2} C_{\text{sign}}(\ell) \right], \quad (1)$$

where $C(\ell)$ is the autocorrelation function of the original signal, C_{sign} is the autocorrelation function of the sign series and ℓ is the lag, considering that the autocorrelation function of linear Gaussian noises follows a power-law, we find an analytic relation between the exponent of correlation of the series, α_{out} ² and the exponent of correlation of the sign series, α_{sign} :

$$\alpha_{\text{sign}} \simeq \alpha_{\text{out}} - \frac{\log(\pi/2)}{2 \log \ell}. \quad (2)$$

We can see that α_{sign} tends asymptotically to α_{out} and the small differences we see in Fig. 2.3 are because of finite size effects. Finally, by means of the decomposition method we can generate binary sequences with the same exponent of correlation that the original series. At last, it is also worth-mentioning that given the fact that in this interval correlations are those in the sign, no matter the correlations in the magnitude. This implies that the coupling between magnitude and sign does not seem to play an important role.

- (iii) $1 \leq \alpha_{\text{in}} \leq 2$. The correlations in the magnitude and the sign increase as correlations in the original series do. Now, α_{mag} (exponent of correlation of the magnitude series) tends asymptotically to α_{out} , i.e. the magnitude controls correlations in this region. On the other hand, the correlations in the sign are given by:

$$\alpha_{\text{sign}} = \frac{1}{2}(1 + \alpha_{\text{in}}) \quad (3)$$

This behavior has also been previously explained by means of the first passage time distributions (Carretero-Campos et al., 2012). We can also deduce from this result

² α_{out} is the output exponent of DFA and its value practically coincides with α_{in}

that $\alpha_{\text{sign}} \leq 3/2$ if $\alpha_{\text{in}} < 2$ (see Appendix A for a proof).

The dependence of α_{sign} on α_{in} (results obtained with DFA) can be summarized as follows:

$$\alpha_{\text{sign}} = \begin{cases} \frac{1}{2} & \alpha_{\text{in}} < \frac{1}{2} \\ \alpha_{\text{in}} & \frac{1}{2} \leq \alpha_{\text{in}} < 1 \\ \frac{1}{2}(1 + \alpha_{\text{in}}) & 1 \leq \alpha_{\text{in}} < 2 \\ \frac{3}{2} & 2 \leq \alpha_{\text{in}} \end{cases} \quad (4)$$

For the magnitude, the asymptotic behavior consists of an uncorrelated zone for $\alpha_{\text{in}} > 3/4$ (in Chapter 3 we prove this is a spurious result of DFA (Carpena et al., 2017)), a transition regime for $3/4 < \alpha_{\text{in}} < 5/4$ and a region where $\alpha_{\text{mag}} \simeq \alpha_{\text{in}}$. This result is in agreement with those shown in (Kalisky et al., 2005) for the series Δx_i^2 (in Chapter 3 we will prove that both series possess similar scaling properties).

Composition of magnitude and sign series

Here, we study systematically the correlations properties of series generated by the composition of independent magnitude and sign series. This procedure can be summarized as follows: first, we generate a fGn with exponent α_{in1} and calculate its magnitude time series, $x_{\text{mag}}(i)$. Then, we generate a fGn with exponent α_{in2} and calculate its sign series $\alpha_{\text{sign}}(i)$. Finally, the composed series, $x_{\text{comp}}(i)$ is obtained: $x_{\text{comp}}(i) = x_{\text{mag}}(i) \cdot x_{\text{sign}}(i)$ (this method was originally proposed in (Kalisky et al., 2005)). Then, we study the correlation properties of the composed series with $\alpha_{\text{in1}}, \alpha_{\text{in2}}$ in the interval $[0.5, 2]$ ($\alpha_{\text{mag}} \in [0.5, 2]$ and $\alpha_{\text{sign}} \in [0.5, 1.5]$). We do not explore the region $\alpha_{\text{in1}}, \alpha_{\text{in2}} > 0.5$ because, as shown previously, in this region the magnitude and sign series are uncorrelated (measured by DFA).

Case $\alpha_{\text{sign}} > 1$.

Here, the correlations in the composed signal are controlled by those in the sign series. In Appendix B we show that the autocorrelation function of the composed series obtained by means of the product of independent magnitude and sign is given by:

$$C(\ell) = C_{\text{sign}}(\ell) \frac{(\pi - 2)C_{\text{mag}}(\ell) + 2}{\pi}, \quad (5)$$

where $C(\ell)$, $C_{\text{mag}}(\ell)$ and $C_{\text{sign}}(\ell)$ are respectively the autocorrelation functions of the composed, magnitude and sign series. Here, we can distinguish two different cases:

- (i) $\alpha_{\text{mag}} < 1$. If the composed series has correlations decaying as a power-law it can be shown from Eq. 5 that the autocorrelation functions of the composed and sign series have similar scaling properties:

$$C(\ell) \sim C_{\text{sign}}(\ell) \quad (6)$$

- (ii) $\alpha_{\text{mag}} \geq 1$. Now, $C_{\text{mag}}(\ell)$ remains constant and we can obtain from Eq. 5 (for long enough scales):

$$C(\ell) \propto C_{\text{sign}}(\ell) \quad (7)$$

We remark that while Eq. 6 is an approximation valid only for long enough scales ℓ , Eq. 7 holds always (see Fig. 2.6 for an example of both situations).

Case $\alpha_{\text{mag}} < 1, \alpha_{\text{sign}} > 1$.

In this case we observe different behaviors when considering short and long scales. At short scales the exponent of correlation of the composed series behaves as $\alpha_1 \simeq \alpha_{\text{mag}}$ whilst at long scales, the exponent of the composed series is $\alpha_2 \simeq \alpha_{\text{sign}}$ (see Fig. 2.7 to visualize examples). From an intuitive point of view, the appearance of this crossover³ can be explained as follows: considering $\alpha_{\text{sign}} > 1$ is unlikely that sign changes its value (from +1 to -1 or the opposite) in an small window (short scales). Then, the short term correlations depend only on the magnitude series. However, at long enough scales, the changes in the sign series will create fluctuations much higher than those created by the magnitude series, and thus, the correlations in the composed series will be controlled by the sign series. Here, we find that the position where the crossover is located, ℓ_c , depends on the size of the series, N , and can be determined analytically, obtaining $\ell_c \propto N^k$, being $k = (\alpha_{\text{sign}} - 1)/(\alpha_{\text{sign}} - \alpha_{\text{mag}})$ (see Fig. 2.8).

Case $\alpha_{\text{mag}} > 1, \alpha_{\text{sign}} > 1$.

Here, the fluctuation $F(\ell)$ might also present a crossover, it will be difficult to observe in practice though. In this case, we obtain analytically that the position of the crossover is proportional to the size of the time series, i.e. $\ell_c \propto N$. This implies the position where the crossover is located grows as the size of the sequence does, thus, making difficult a clear observation of this phenomenon. Here, we can distinguish two different regions:

³The term *crossover* denotes a change in the exponent of correlation of the time series. Looking at the $\log(F(\ell))$ vs. $\log(\ell)$ plots it implies a change in the slope.

- (i) $\alpha_{\text{mag}} \geq \alpha_{\text{sign}}$. In this case, in all experiments we obtain $\ell_c > N$, so it is not possible to visualize the crossover (see Fig. 2.9).
- (ii) $\alpha_{\text{mag}} > \alpha_{\text{sign}}$. Here, in some situations we obtain $\ell_c > N$, although the values for the fluctuation of the magnitude and sign are quite similar, thus, making difficult the visualization of the crossover (see Fig. 2.9). Moreover, when we apply DFA we only analyze scales up to $\ell = N/10$ in order to have a big enough number of windows (Hu, 2001).

Definitely, in this region we barely observe crossover and the composed series will have one single scaling, with value $\alpha = \min\{\alpha_{\text{mag}}, \alpha_{\text{sign}}\}$ (see text in Chapter 2 for a proof of this result). Another conclusion we can draw from this case is the fact that it is not possible to generate synthetic series with exponent of correlation $\alpha > 1.5$ (in average). Here, we calculate the exponent of correlation of composed series obtained with $\alpha_{\text{mag}} = 1.9$ and $\alpha_{\text{sign}} = 1.5$ ($\alpha_{\text{in1}} = 2.0$, $\alpha_{\text{in2}} = 2.0$). In Fig. 2.10 we see the distribution obtained from the exponent of correlation of the composed series. We see a distribution with a sharp peak in $\alpha = 1.5$ ($\min\{\alpha_{\text{mag}}, \alpha_{\text{sign}}\}$). There are some time series where the exponent of correlation is $\alpha > 1.5$ due to the fact that in some cases the scaling of the composed series is not good, together with the cases where the crossover is observable.

Lastly, we summarize the results obtained for composition in the next table:

α_{sign}	α_{mag}	α	CROSSOVER
< 1	$[0.5, 2]$	α_{sign}	no
> 1	< 1	$\ell < \ell_c \quad \alpha_1 = \alpha_{\text{mag}}$	$\ell_c \propto N^k$ $k = \frac{\alpha_{\text{sign}} - 1}{\alpha_{\text{sign}} - \alpha_{\text{mag}}}$
		$\ell > \ell_c \quad \alpha_2 = \alpha_{\text{sign}}$	
> 1	> 1	$\min\{\alpha_{\text{mag}}, \alpha_{\text{sign}}\}$	not observable

Tabla 1: Results obtained for composition by means of independent magnitude and sign time series.

Multifractal properties of composed series

Up to now, we have focused only on the linear correlations. However, it is well-known that time series with correlations in the magnitude also possess nonlinear and multifractal properties (Ashkenazy, 2003; Kalisky et al., 2005) (In Chapter 3 we will prove that this does not hold always). In this section, we study the multifractal properties of composed series by means of independent magnitudes and signs. Considering that in the previous section we proved the existence of crossovers, this behavior could imply the existence of

two different multifractal spectra. Then, we will study only the region $\alpha_{\text{sign}} < 1$, where the composed series has one single scaling and the correlations are controlled by those in the sign ($\alpha = \alpha_{\text{sign}}$). We should also take into account that during simulations numerical instabilities can appear when $\alpha_{\text{mag}} > 1.2$. Then, we finally study the nonlinear properties of composed series in the region $\alpha_{\text{mag}} \in [0.5, 1.2]$ and $\alpha_{\text{sign}} \in [0.5, 1]$. Specifically, we calculate the multifractal spectrum by means of MFDFA (see Chapter 1) for each composed series and study systematically two properties: the width of the spectrum, $\Delta\zeta$, and the location of the center of the spectrum, ζ_{max} . We observe that the width of the spectrum only depends on the exponent of correlation of the magnitude, α_{mag} , being this dependence almost linear (see Fig. 2.11). Consequently, the width of the spectrum is negligible for values $\alpha < 0.75$ and starts to grow as the magnitude is correlated ($\alpha > 0.75$). Moreover, the exponent of correlation of the sign, α_{sign} does not affect the width of the spectrum, but it controls its location (see Fig. 2.11) (there is also a linear dependence between the location of the center of the spectrum and the exponent of correlation of the sign). These properties let us create a multifractal generation model which allows us to control both linear and nonlinear properties. In fact, we can control the linear correlations of the signal with α_{sign} and the strength of nonlinearities with α_{mag} .

Autocorrelation function of magnitude and sign in fractional Gaussian noises

Up to now, we have studied the correlation properties of fGns and composed series by means of DFA. However, DFA is a method which measures correlations indirectly and only offers an accurate value of the exponent of correlation when the fluctuation follows a power-law dependence with the scale of observation (also the power spectrum has a power-law dependence with the frequency (Rangarajan and Ding, 2000)). Nevertheless, in most cases when analyzing real data, there is no power-law dependence between fluctuation and observation scale. This is the reason why we focus on the autocorrelation function, which despite being noisy and sensitive to the size of the sequence, can be of great utility to better understand the properties of the time series. Given a fGn $\{x_i\}$, we first obtain the analytic expression of the autocorrelation function of the magnitude $C_{|x|}$ as a function of its autocorrelation C_x . We also obtain analytic expressions relating the autocorrelation of the sign series C_s with the autocorrelation C_x , and the autocorrelation function of quadratic series C_{x^2} with C_x . Finally, we propose a new index of nonlinearity based on the deviation of the autocorrelation function of the magnitude in nonlinear series respect to the autocorrelation expected in the case of linear Gaussian noises.

Autocorrelation function of magnitude in fractional Gaussian noises

Given a fGn $\{x_i\}$ with autocorrelation function $C_x(\ell)$, we obtain the autocorrelation of its magnitude series ($\{|x_i|\}$) $C_{|x|}(\ell)$ as a function of $C_x(\ell)$. Considering two random variables $\{X, Y\}$, both with zero mean and unit standard deviation and the bivariate Gaussian distribution (Johnson et al., 2002):

$$\rho(x, y) \equiv \text{Prob} \{X = x, Y = y\} = \frac{1}{2\pi\sqrt{1-k^2(\ell)}} \exp \left\{ -\frac{x^2 + y^2 - 2k(\ell)xy}{2[1-k^2(\ell)]} \right\}, \quad (8)$$

where $k = \langle xy \rangle$ is the covariance of variables X and Y , which also coincides with their correlation taking into account that both of them have zero mean and unit standard deviation. The covariance of $|X|$ and $|Y|$ is given by:

$$k_{\text{mag}} \equiv \langle |x||y| \rangle - \langle |x| \rangle \langle |y| \rangle = \int_{-\infty}^{\infty} |x| dx \int_{-\infty}^{\infty} |y| dy \rho(x, y) - \frac{2}{\pi}, \quad (9)$$

where we have used that $\langle |x| \rangle = \langle |y| \rangle = \sqrt{2/\pi}$. Finally, given that $X \equiv x_i$ and $Y \equiv x_{i+\ell}$ the autocorrelation of the magnitude series can be expressed as:

$$C_{|x|}(\ell) = \frac{2 \left[C_x(\ell) \sin^{-1}(C_x(\ell)) - 1 + \sqrt{1 - C_x^2(\ell)} \right]}{\pi - 2}. \quad (10)$$

$C_{|x|}(\ell)$ always takes positive values in Eq. 10, thus, the magnitude cannot be anticorrelated (see Fig. 3.1). We also see that despite being the magnitude series correlated, fGns are linear (experimental points perfectly fits Eq. 10), then, the fact that a time series possesses correlations in its magnitude series, does not necessary implies that the series has nonlinear properties. Furthermore, considering the first non-vanishing term of the Taylor expansion for small values in the autocorrelation of the magnitude we obtain:

$$C_{|x|}(\ell) \simeq \frac{1}{\pi - 2} C_x^2(\ell), \quad (11)$$

thus, the autocorrelation of the magnitude behaves as the square of the autocorrelation of the original signal for small values (see Fig. 3.1). Taking into account that Eq. 1.12 is valid for fractional Gaussian noises with $0 < H \leq 1$ where H is the Hurst exponent we can draw some conclusions. Having in mind that $H = (2 - \gamma)/2$ where γ is the autocorrelation exponent (see Chapter 1 for a detailed description of different exponents of correlation) we note that the time series will possess correlations when $\gamma < 1$ ($H > 0.5$). On the contrary,

when $\gamma > 1$ ($H < 0.5$) the time series possesses negative correlations or anticorrelations. Then, considering Eq. 11, we obtain that the autocorrelation of the magnitude decays as:

$$C_{|x|}(\ell) \propto \frac{1}{\ell^{2\gamma}}, \quad (12)$$

taking place two different scenarios:

- (i) $2\gamma < 1 \Rightarrow H > 3/4$. The correlations in the magnitude decays slower than $1/\ell$, leading to long-range correlations in the magnitude.
- (ii) $2\gamma > 1 \Rightarrow H < 3/4$. In this case, although the correlations are still positive (see Fig. 3.1) they decay faster than $1/\ell$, reaching the noise threshold quickly. In fact, Fluctuation Analysis Techniques (FA and DFA) are not able to detect correlations in the magnitude series when $H < 3/4$ (Carpena et al., 2017) (DFA measures no correlations, $\alpha_{\text{mag}} = 0.5$). Then, the results obtained for correlations in the magnitude by means of DFA when $H < 3/4$ ($H = \alpha$ in the interval $H \in [0, 1]$) in Chapter 2 are spurious and the time series presents short term but positive correlations which DFA is not able to detect (see Fig. 3.2).

Autocorrelation function of quadratic series

It is widely known that scaling properties of magnitude series $\{|x_i|\}$ are similar to those in the quadratic series $\{x_i^2\}$ (Kalisky et al., 2005). Due to this, and for the sake of simplicity, sometimes the autocorrelation of the quadratic series is studied instead of the autocorrelation of the magnitude. Here, we obtain the analytic relation for fGns $\mathcal{N}(0, 1)$ between the autocorrelation of the quadratic series C_{x^2} and the autocorrelation of the original series C_x . C_{x^2} is given by:

$$C_{x^2}(\ell) = \frac{\langle x_i^2 \cdot x_{i+\ell}^2 \rangle - \langle x_i^2 \rangle \langle x_{i+\ell}^2 \rangle}{\sigma_{x^2}^2}, \quad (13)$$

and the covariance of $\{x_i^2\}$, considering $x_i = x$, $x_{i+\ell} = y$:

$$K_{sq}(\ell) = \langle x^2 \cdot y^2 \rangle - \langle x^2 \rangle \langle y^2 \rangle, \quad (14)$$

Expressing the covariance as an integral:

$$K_{sq}(\ell) = \int_{-\infty}^{\infty} x^2 dx \int_{-\infty}^{\infty} y^2 dy \cdot \rho(x, y) - 1 = 2k_x^2(\ell) \quad (15)$$

being $k_x(\ell)$ the covariance of $\{x_i\}$, ρ the bivariate Gaussian distribution (Johnson et al., 2002) and considering $\langle x^2 \rangle = \langle y^2 \rangle = 1$. Finally, having in mind that our fGns have zero mean and unit standard deviation we obtain:

$$C_{x^2}(\ell) = C_x^2(\ell) \quad (16)$$

This result implies that the autocorrelation of the quadratic series C_{x^2} cannot take negative values (see Fig. 3.3). Furthermore, this result fully explains the reason why the magnitude and quadratic series present similar scaling properties (Kalisky et al., 2005) as for long enough values of ℓ ($C_x \ll 1$), C_x and $C_{|x|}$ take almost identical values. Another way to prove this result consists of comparing Eq. 11 and Eq. 16 where we can see that both present the same scaling exponent.

Autocorrelation function of sign series

Sign series also provides with important information about the system under study. In fact, it is instrumental when studying the first-passage time distributions (Carretero-Campos et al., 2012) or keyword detection in texts (Carpena et al., 2009; Ortuno et al., 2002). There is an analytic relation involving the autocorrelation of the sign C_s and the autocorrelation of the original signal C_x for Gaussian series, which was originally proposed by (Apostolov et al., 2008):

$$C_s = \frac{2}{\pi} \arcsin(C_x). \quad (17)$$

Here, he prove this relation by two different ways (see Chapter 3) and, moreover, we prove that this relation is also valid for non-Gaussian series. Considering the first non-vanishing term of the Taylor expansion of Eq.17 we obtain (see Fig. 3.4):

$$C_s = \frac{2}{\pi} C_x \quad (18)$$

Taking into account that both original and sign series decays as a power-law with the same exponent of correlation:

$$C_s(\ell) \propto \frac{1}{\ell^\gamma}, \quad (19)$$

leading to two interesting cases (see Fig. 3.5):

- (i) $\gamma < 1 \Rightarrow \alpha_{\text{sign}} > 0.5$. Correlations in the sign decay slower than $1/\ell$, thus, there are long-range correlations. Moreover, the exponent of correlation of the sign series is identical to that in the original series.
- (ii) $\gamma > 1 \Rightarrow \alpha_{\text{sign}} < 0.5$. Correlations in the sign decay faster than $1/\ell$, thus, the sign

series will present negative and short-term correlations (see Eq. 3.100). Considering Eq. 18, the original and sign series must always have the same correlation exponent. However, when decomposing fGns (see Fig. 2.3) and measured correlations with DFA we obtained $\alpha_{\text{sign}} = 0.5$ for anticorrelated series. In fact, this is another spurious result of DFA which is not able to detect short-term correlations in sign series (Carpena et al., 2017)

Finally, combining Eqs. 10 and 17 we can obtain the analytic relation between the auto-correlation of magnitude and sign in fGns (see Fig.3.6):

$$C_{|x|} = \frac{2}{\pi - 2} \left[\frac{\pi}{2} C_s \sin \left(\frac{\pi}{2} C_s \right) + \cos \left(\frac{\pi}{2} C_s \right) - 1 \right], \quad (20)$$

and its Taylor expansion for small values of ℓ ;

$$C_{|x|} = \frac{\pi^2}{4(\pi - 2)} C_s^2 + \mathcal{O}(C_s^4) \quad (21)$$

Nonlinear model

Up to now, we have studied the analytic relations involving the autocorrelation functions of original, magnitude and sign series (Eqs. 10,17,20). However, these equations are only valid for linear Gaussian noises, and, in the case of nonlinear noises, the deviation from these equations can be used as an index of nonlinearity. Moreover, this index of nonlinearity does not require the presence of scaling in the time series under study, which can be a great advantage over the standard techniques used to measure nonlinearity (DFA and Multifractal DFA). We focus on Eq. 10, because it is widely known that correlations in the magnitude are linked to nonlinear properties in the series under study (Ashkenazy, 2001; Kalisky et al., 2005) (although we have seen that it is possible that time series with correlated magnitude do not have nonlinear properties). To show the effect of nonlinearity we generate nonlinear series by means of composition of independent magnitudes and signs (Kalisky et al., 2005). The autocorrelation function of these kind of series is given by (Gómez-Extremera et al., 2016):

$$C_x(\ell) = C_s(\ell) \frac{(\pi - 2)C_{|x|}(\ell) + 2}{\pi} \quad (22)$$

where $C_{|x|}$ is the auto-correlation of the magnitude series, generated from a fGn with Hurst exponent H_1 and C_s is the auto-correlation of the sign series, generated from a fGn with Hurst exponent H_2 . Then, we can write the expression of the autocorrelation

function of the composed series in terms of H_1 and H_2 :

$$C_x(\ell) \simeq \frac{2H_2(2H_2 - 1)}{\pi^2 \ell^{2-2H_2}} \cdot \left[\frac{H_1^2(2H_1 - 1)^2}{\ell^{4-4H_1}} + 2 \right]. \quad (23)$$

In Fig. 3.7 we represent several examples of nonlinear Gaussian noises. In all cases the noises have been generated by means of sign series with $H_2 = 0.85$, thus, maintaining the linear correlations in the series (controlled by the sign series). We also observe that curves are different as we modify the correlations in the magnitude. Then, given a time series $\{x(i)\}$ we propose a new measure of nonlinearity based on the deviation of $C_{|x|}$ respect to the expected value in case of linear Gaussian noises $C_{|x|,\text{linear}}$. Then, the index of nonlinearity, Δ is defined as follows:

$$\Delta = \sum_{\ell=1}^{\ell_{\max}} [C_{|x|}(\ell) - C_{|x|,\text{linear}}(\ell)]^2, \quad (24)$$

where $C_{|x|}$ is the autocorrelation function of the magnitude of the series under study, $C_{|x|,\text{linear}}$ is the expected value in the case of a linear Gaussian noise and ℓ_{\max} is the value of ℓ up to we compute the autocorrelation function. In the case of nonlinear series generated by means of composition we can also obtain analytic expressions for the nonlinearity index Δ and we observe in Fig. 3.8 that Δ increases substantially for values $H_1 > 0.75$ ($\alpha_{\text{mag}} > 0.5$). This results agrees with those obtained when analyzing the multifractal properties of composed series by means of Multifractal DFA. In fact, for both index (Δ and the width of the spectrum $\Delta\zeta$) the multifractal and nonlinear properties increase once the magnitude is correlated. Moreover, for values $H_1 < 0.75$ (range where DFA is not able to detect correlations in the magnitude) the composed series also present nonlinearities, although both parameters (the width of the multifractal spectrum and the nonlinearity index) take very small values.

Measures of nonlinearity in real time series

In the field of time series, the concept of nonlinearity can be interpreted from different points of view. One intuitive definition consists on considering as nonlinear series those generated in systems ruled by nonlinear dynamics, i.e. the values of the time series depend on nonlinear expressions such as quadratic, logarithms, trigonometric, exponential, etc. However, in most of the cases, we do not know anything about these equations and the goal is just to find them. Another definition of nonlinearity more suitable for practical purposes is found in (Schreiber and Schmitz, 2000). Nonlinear series are those whose

Fourier phases are not random. Then the presence of nonlinearities in a time series can be assessed as follows: (i) Given a time series, compute its Fourier transform, randomize its Fourier phases, and transform it back. The resulting surrogate series preserves the distribution of the data and the linear correlations because its power spectrum remains unchanged (Schreiber and Schmitz, 2000). Then, some statistics is evaluated between the original and surrogate series and if there are significant differences, it means that the original Fourier phases were not random and the null hypothesis of linearity can be rejected. Sometimes, instead of accepting or rejecting the null hypothesis the goal is to compare the degree of nonlinearity of two different series e.g. records obtained under different physiological conditions, and the value of the statistics (Δ in our case) is directly used as an index of nonlinearity.

Here we will analyze the heartbeat fluctuation time series. From the pioneering papers (Peng et al., 1993b) much attention has been paid to RR signal (RR denotes the period between two consecutive heartbeats). In fact, it has been shown that correlations in RR series are of great utility to detect alterations due to pathologies or age (Goldberger et al., 2002; Ivanov et al., 1996), different physiological states (Ivanov et al., 1999) or different state of fitness (Aubert et al., 2003). In most of the cases, papers are only focused on linear correlations, being the nonlinear correlations instrumental though (their absence has been linked to age and pathological situations (Ivanov, 1999a)).

Nonlinearity during rest and exercise

It is well-known that heart behaves totally different during rest and exercise. The most noticeable change is the abrupt increase of the cardiac rhythm and reduction of heart rate variability (HRV) once the subject starts to do exercise (Sarmiento et al., 2013). Besides these changes, it has also been shown that exercise modifies the power spectrum, decreasing low-frequency components (Anosov et al., 2000; Brodie, 2006; Sarmiento et al., 2013) and introducing high-frequency components clearly affected by the respiratory rhythm (Lewis and Short, 2010; Martín-González and García-Manso, 2012). With respect to the linear correlations the short-term exponent of the DFA (α_1) is commonly used (despite the usual lack of scaling in RR series and being strongly affected by respiration (Perakakis et al., 2009)). It can be found that short-term correlation increases when exercising (Karasik et al., 2002; Platasa et al., 2008) and also when increasing the intensity of the exercise (A. J. Hautala et al., 2003), although the opposite result can also be found (Tulppo et al., 2001). Summarizing, in spite of these contradictory results obtained with α_1 it seems that exercise reduces the complexity of RR series due to the breaking of the balance between sympathetic and parasympathetic branches of the Autonomic Nervous System (NSA).

Here, we obtain nonlinear properties from recordings of amateur and professional players during rest and exercise (running at moderate pace for amateur players and stress test for professional players). We compute the nonlinearity index Δ as follows: (i) we obtain the increment series ΔRR (it has stationarity, at least in the weak sense and its probability distribution has certain similarity with the Gaussian distribution) and (ii) we transform the distribution of ΔRR to a Gaussian $\mathcal{N}(0, 1)$ and obtain $\{\Delta RR_{\mathcal{N}}\}$ (see Sec. 1.8). Then we compute the nonlinearity index of $\{\Delta RR_{\mathcal{N}}\}$ by means of Eq. 24 and using $\ell_{\max} = 10$.

In the case of amateur players we have a database of 10 recordings (rest and running), all of them male (age 23.0 ± 2.1 yr without cardiac pathologies, see Table 4.1). Results in Fig. 4.3 show a clear difference of nonlinearity between rest and exercise⁴ $p = 3 \cdot 10^{-4}$.

Following the same procedure previously mentioned we also compare the nonlinearity index in professional soccer players. We use 12 recordings, all of them from male without heart pathology (age 22.1 ± 3.4) (see Table 4.2) during rest and stress test. Again, results in Fig. 4.5 show a clear difference between rest and stress test ($p = 8 \cdot 10^{-6}$).

Lastly we compare results obtained from amateur and professional players and observe that nonlinearity is higher in rest for professional players $p = 0.047$ (we associate this result to the fact that professional players are in better state of fitness than amateur players). However, during exercise, we obtain a higher nonlinearity index for amateur players (they were running at a moderate pace whilst professional players were doing an stress test) ($p = 0.17$, thus, although we notice a difference between mean values, we cannot consider this difference as significant). A possible explanation for this phenomenon is that the intensity of the exercise is linked to the nonlinear properties of the time series, i.e the time series loses complexity as the intensity of the exercise increases.

Nonlinearity during rest: aerobic and non-aerobic training

There are some papers in the literature which show the permanent effects of exercise in the heart dynamics. For example, studying correlations at different levels of intensity (A. J. Hautala et al., 2003), linear correlations during rest and exercise (Karasik et al., 2002), linear correlations after and before a long period training (Tulppo et al., 2003) and HRV study in some situations related to physical exercise (Bernardi et al., 1996; Byrne et al., 1996; Davy et al., 1997; Goldsmith et al., 1992; Levy et al., 1998; Seals and Chase, 1989). Nevertheless, very few papers study the influence of exercise in the complexity of the signals. In fact, the very few papers published are contradictory at a certain way.

⁴Due to the fact that during exercise time series are longer than in rest, we split our exercise series in sub-series with the same size that their corresponding rest series, in order to avoid finite size effects.

The only point of agreement is that periodic exercise decreases the cardiac rhythm and increases HRV (during rest).

Here we compare the nonlinearity index during rest in two groups of people who follow totally different training methods. On the one hand we have a database with recordings from 28 soccer players (amateur and professional), all males with age 23.0 ± 4.1 yr (see Table 4.3) who always have aerobic training during sessions. On the other hand we have a 34 recording database from amateur body-builders, age 28.0 ± 6.1 yr (see Table 4.4) who do not include any cardiovascular exercise (aerobic) in their sessions. When comparing the nonlinearity index (calculated in the same way we did in previous cases) we obtain that nonlinearity during rest is higher for soccer players, observing clear group difference ($p = 7.6 \cdot 10^{-4}$). This result suggests that aerobic training makes nonlinearity increase.

Other preliminary results

Here, we present some other preliminary results we have obtained with heartbeat time series. Specifically, we know that, given a RR recording which has a rest stage, exercise and rest after exercise (recovering), the behavior of the rest stages (before and after exercise) are completely different. In fact, in Fig. 4.9 we show a register from the author of this document in which appear a rest period (approximately 10 min), 50 minutes running at 15 km/h and a rest period (recovering after exercise, approximately 10 min). Having a look at the figure it is clear that the initial rest period (before exercise) has completely different statistical properties than the recovering period (much lower mean value and much less stationary than the initial rest period). In fact, the recovering period has been used in different ways to reveal important information: the number of heartbeats recovered per minute is an index of state of fitness⁵, as a predictor of mortality (Cole et al., 1999), to assess pathologies (Imai et al., 1994; Lipinski et al., 2004), analyzing its relation with the NSA (Savin et al., 1982), etc.

Then, despite present significant differences respect to the rest before exercise, we hypothesize that there will also exist significant differences between the recovery period and exercise. To do so, we analyze a database from Prof. Bernaola-Galván in which registers contain a running period (approximately one hour) and then, a recovery period (around 10 min). We show the results in Fig. 4.10, where we obtain that the nonlinearity index Δ is much higher (in average) during recovery ($p = 6 \cdot 10^{-4}$). This result is interesting for the future because: (i) in spite of the high nonstationarity of the RR series during exercise and recovering period, our measure is able to detect that the subject is recovering (a

⁵Given two subjects exercising at the same intensity and duration (one in a great state of fitness and the other sedentary) the sedentary will spend much more time to recover the mean RR in rest.

significant difference exists) and (ii) we could design an experiment in which subjects in different state of fitness did exercise following the same protocol (activity, duration, intensity, etc) we could correlate the nonlinearity index with some other statistical properties of the signal and then, establish a new way to assess the state of fitness of the subject, detect possible pathologies, etc.

Introducción

Un gran número de fenómenos asociados a distintos campos, entre los que podemos resaltar Física, Fisiología, Economía, Biología, Geología, Meteorología, etc, ofrecen señales⁶ de salida que aunque en apariencia puedan aparecer erráticas y ruidosas, poseen correlaciones de largo alcance y estructura fractal (invariancia frente a cambios de escala) (Mandelbrot, 1975). Algunos ejemplos destacados donde se han detectado correlaciones de largo alcance son: el ADN humano (Hackenberg et al., 2011, 2012; Li and Kaneko, 1992; Peng et al., 1992; Voss, 1992), la música (Voss and Clarke, 1978), Hidrología (Hurst, 1951), corazón humano (Peng et al., 1993b), búsqueda de palabras clave en textos (Carpena et al., 2009; Ortuno et al., 2002), economía (Peters, 1994), etc. Además, en algunos casos, el valor de la correlación está íntimamente ligado al estado del sistema, e.g. en el corazón el valor de la correlación cambia drásticamente al analizar series pertenecientes a sujetos sanos o enfermos (Ivanov, 1999a), en diferentes estados fisiológicos (Ivanov et al., 1999) o en función de la edad (Iyengar et al., 1996).

Sin embargo, también existen sistemas donde series con las mismas correlaciones de largo alcance presentan propiedades dinámicas totalmente diferentes (Ashkenazy, 2001). Estos sistemas presentan interacciones complejas (cuadráticas, logarítmicas, exponenciales, trigonométricas, etc) que van mucho más allá de la linealidad, por lo que un estudio exhaustivo de las mismas puede arrojar información complementaria muy importante para comprender la dinámica del sistema. En la práctica es muy difícil averiguar la dependencia funcional de estas interacciones y el logro, en la mayoría de los casos, consiste en evaluar el grado de complejidad o no-linealidad del sistema (Ivanov, 1999a; Kalisky et al., 2005). Una forma práctica de romper esta degeneración consiste en estudiar de forma separada las correlaciones de la serie módulo (valor absoluto de la serie) y las de la serie signo (+1 si la serie es positiva y -1 si la serie es negativa) ya que: (i) las correlaciones de la serie módulo han sido asociadas a propiedades no lineales y multifractales del sistema (Ashkenazy, 2001, 2003; Gómez-Extremera et al., 2016; Kalisky et al., 2005) y (ii) las correlaciones de la serie signo controlan las correlaciones lineales de la serie de la que

⁶A lo largo de este trabajo se utilizarán los términos señal y serie temporal de forma indistinta

procede (Ashkenazy, 2001, 2003; Carretero-Campos et al., 2012). Por tanto, el estudio de las propiedades de correlación de las series módulo y signo puede ser de gran utilidad a la hora de comprender los mecanismos de acoplamiento entre ambas y permite arrojar información sobre sistemas complejos donde la dinámica está controlada por dos mecanismos competidores, e.g. los aumentos y descensos del ritmo cardíaco están controlados por las ramas simpática y parasimpática del Sistema Nervioso Autónomo respectivamente.

El DFA (Detrended Fluctuation Analysis) (Peng et al., 1994b) (ver Sección 1.3.2) es la técnica por excelencia utilizada en la bibliografía para la estimación de correlaciones. Esta técnica elimina tendencias en series temporales y es capaz de medir (de forma indirecta) el exponente de correlación en series cuyas correlaciones decaen en forma de ley de potencias⁷. Sin embargo, el DFA presenta un importante inconveniente: cuando se aplica a datos reales, en la mayoría de los casos, la fluctuación no tiene una dependencia funcional en forma de ley de potencias con la escala de observación, por lo que su uso puede conducir a resultados espurios (aún no está claro qué significan las desviaciones sobre la ley de potencias). Además, es necesario que la serie bajo estudio sea suficientemente larga para conseguir buena estadística. Para el módulo y el signo, incluso en los casos donde estos requisitos se cumplen, el DFA no es capaz de estimar de forma precisa sus correlaciones en algunos casos (Carpena et al., 2017). Para afrontar este problema, una interesante opción es el estudio de la función de autocorrelación, que a pesar de ser ruidosa y sensible al tamaño de la serie, nos facilita la información que estamos buscando.

A raíz de lo expuesto, en este trabajo nos proponemos estudiar las propiedades de series módulo y signo mediante técnicas de análisis de fluctuación (DFA) y función de autocorrelación, detectar los posibles resultados espurios ofrecidos por el DFA y basándonos en el hecho contrastado de que las correlaciones del módulo están asociadas a las propiedades no lineales de la serie de la que procede (más adelante veremos que esto no es siempre así), establecer una nueva medida de no-linealidad a partir de la función de autocorrelación. Finalmente, estudiaremos propiedades no lineales de la serie de latidos cardíacos en diferentes condiciones y estados fisiológicos. Para cumplir estos objetivos, este trabajo se estructura de la siguiente forma:

- En el capítulo 1 estudiamos conceptos básicos de series temporales, modelos que reproducen series presentes en la naturaleza, así como un amplio abanico de herramientas y técnicas utilizadas en el estudio de series temporales.
- En el capítulo 2 estudiamos mediante el uso del DFA las correlaciones lineales de series módulo y signo procedentes de series lineales Gaussianas. También estudiamos

⁷Cuando una serie tiene correlaciones que decaen en forma de ley de potencias, su fluctuación también tiene una dependencia funcional en forma de ley de potencias con la escala de observación.

las correlaciones de series obtenidas a través de series módulo y signo que proceden de diferentes series para entender los mecanismos de acoplamiento entre el módulo y el signo. Por último, estudiamos propiedades no-lineales y multifractales de dichas series mediante el uso del Multifractal Detrended Fluctuation Analysis (una versión modificada del DFA para el estudio de no-linealidad).

- En el capítulo 3 obtenemos las relaciones analíticas de la función de autocorrelación del módulo en ruidos lineales Gaussianos frente a la función de autocorrelación de la propia señal. En el caso del signo, ya existía una relación analítica que relaciona su función de autocorrelación con la de la señal de la que procede en el caso de series con distribución de valores Gaussianas (Apostolov et al., 2008). Probamos que esta relación no es solo válida para series Gaussianas, sino que es válida para cualquier serie con distribución de valores simétrica. Además, comparamos los resultados obtenidos, tanto para el módulo como para el signo, con los obtenidos en el capítulo 2 mediante el uso del DFA para detectar posibles resultados espurios. Por último, estudiamos la función de autocorrelación del módulo en las series obtenidas a través del producto de módulos y signos que proceden de series diferentes, y en base a esto, proponemos una nueva medida de no-linealidad basada en la desviación que experimenta la autocorrelación de la serie módulo en una serie no lineal frente a la que cabría esperar en el caso de que la serie fuese lineal.
- En el capítulo 4 aplicamos esta nueva medida de no-linealidad a la serie de latidos cardíacos. En concreto, comparamos el comportamiento del corazón durante reposo y ejercicio suave, reposo y test de esfuerzo y los efectos permanentes del ejercicio ocasionados por el entrenamiento aeróbico frente al entrenamiento de fuerza.
- Finalmente, en el capítulo 5 se presentan las conclusiones del trabajo y las perspectivas de futuro del mismo.

Capítulo 1

Métodos para el estudio y análisis de series temporales

Tradicionalmente se ha considerado que muchos sistemas presentes en la naturaleza, entre los que se incluyen los sistemas fisiológicos, están regidos por el principio de *homeostasis*, por el que el sistema se mantendría en un estado de equilibrio en ausencia de perturbaciones, y en caso de haberlas, el sistema volvería al estado inicial de equilibrio una vez concluida la perturbación y en un periodo de tiempo llamado *tiempo de relajación* (Bernard, 1878; Hyndman, 1974; Van Der Pol and Van der Mark, 1928). Además, estos sistemas presentan una única escala temporal, relacionada con el tiempo de relajación.

En cambio, estudios recientes muestran que dichos sistemas poseen fluctuaciones ruidosas que se manifiestan en múltiples escalas temporales y cuyo comportamiento se asemeja al de sistemas físicos fuera del equilibrio (Bassingthwaighte et al., 1994; Malik and Camm, 1995; Peng et al., 1994a). Analizando en detalle estas fluctuaciones, se observa que presentan correlaciones de largo alcance, es decir, el espectro de potencias de una determinada propiedad del sistema presenta una dependencia funcional con la frecuencia de $1/f^\beta$, con $\beta > 0$, (de ahí que estas señales se denominen *ruido 1/f*) lo que implica una invariancia frente a cambios de escala temporal, o dicho de otra forma, presentan un comportamiento *fractal*, concepto que fue introducido por B. Mandelbrot (Mandelbrot, 1975). Algunos ejemplos destacados donde se ha puesto de manifiesto la existencia del ruido $1/f^\beta$ son el ADN humano (Hackenberg et al., 2011, 2012; Li and Kaneko, 1992; Peng et al., 1992; Voss, 1992), la música (Voss and Clarke, 1978), Hidrología (Hurst, 1951), el corazón (Peng et al., 1993b), búsqueda inteligente de palabras (Carpena et al., 2009; Ortuno et al., 2002), Economía (Peters, 1994), etc. Por tanto, para el estudio y comprensión de la dinámica de estos sistemas son imprescindibles modelos que sean capaces de generar ruidos artificiales cuyo espectro de potencias decaiga en forma de ley de potencias. Uno de los modelos más

utilizados para reproducir el ruido $1/f^\beta$ es el *ruido fraccionario Gaussiano* (fGN) (ver Sec. 1.2).

Este capítulo se organiza de la siguiente forma: En la Sec. 1.1 abordamos algunos conceptos básicos relativos a series temporales, en la Sec. 1.2 estudiamos el origen del ruido fraccionario Gaussiano a partir del trabajo de Hurst, en la Sec. 1.3 estudiamos técnicas empleadas para obtener propiedades de correlación de señales mediante el análisis de fluctuación. En la Sec. 1.4 analizamos el método de Filtrado de Fourier, que es uno de los métodos por excelencia para la creación de fGns de forma sintética. Por último, en las secciones posteriores estudiamos algunas técnicas para el estudio de propiedades multifractales, transformada wavelet, aleatorización de fases y distribución de valores de la serie.

1.1. Serie Temporal

Una serie temporal es una secuencia $\{x_i\}_{i=1,\dots,N}$ de medidas obtenidas a partir de un observable X en sucesivos intervalos de tiempo $t = t_k$. El intervalo de tiempo entre dos medidas consecutivas de X (muestreo) no necesariamente ha de ser igual, i.e. $\exists i, j \mid t_{i+1} - t_i \neq t_{j+1} - t_j$, aunque de ahora en adelante supondremos un muestreo equiespaciado ($\forall i : t_{i+1} - t_i = \text{cte}$) salvo que explícitamente se indique lo contrario.

1.1.1. Medidas de dependencia: covarianza y autocorrelación

En general, la probabilidad de que una variable aleatoria X tome un valor menor o igual que x ($P(X \leq x)$) puede ser expresada en términos de la función de densidad de probabilidad acumulada:

$$F_X = P(X \leq x) = \int_{-\infty}^x p_X(x) dx, \quad (1.1)$$

siendo $p_X(x)$ la función de densidad de probabilidad del suceso X , que satisface que $p_X(x) \geq 0$ y $\int_{-\infty}^{+\infty} p_X(x) dx = 1$. En el caso de N variables aleatorias X_i con $i = 1, 2, \dots, N$, la probabilidad de que X_1, X_2, \dots, X_N sea menor o igual que los valores constantes x_1, x_2, \dots, x_N viene dada por la distribución de probabilidad conjunta (ver (Schumann, 2010)):

$$F(X_1, X_2, \dots, X_N)(x_1, x_2, \dots, x_N) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_N \leq x_N). \quad (1.2)$$

Cuando la probabilidad conjunta en Eq. 1.2 puede expresarse como un producto de las funciones de distribución acumulada de cada una de las variables aleatorias del conjunto X_i , i.e. $F(X_1, X_2, \dots, X_N)(x_1, x_2, \dots, x_N) = \prod_{i=1}^N P(X_i \leq x_i)$, dicho conjunto de variables es *independiente*, y es *dependiente* en la situación contraria. Otro concepto muy importante a tener en cuenta son los momentos de orden n . Las funciones de distribución acumulada de probabilidad $F_{X_i} = P(X_i \leq x)$, y sus derivadas, las densidades de probabilidad, $p_{X_i}(x) = \partial F_{X_i}(x)/\partial x$ facilitan los momentos de orden n de la variable aleatoria X_i , que en general se definen como:

$$\mu_{X_i}^{(n)} = E[X_i^n] = \langle X_i^n \rangle = \int_{-\infty}^{\infty} x^n p_{X_i}(x) dx, \quad n = 1, 2, 3, \dots \quad (1.3)$$

donde $\langle \cdot \rangle$ denota el promedio y E el valor esperado. Algunos órdenes de momentos muy relevantes son el momento de orden 1 y momento de orden 2 que representan la media y la varianza respectivamente.

Con todo lo expuesto, estamos en condiciones de introducir el concepto de covarianza, que es una de las medidas más utilizadas a la hora de establecer el grado de dependencia o independencia entre dos variables aleatorias X_i y X_j ($1 \leq i, j \leq N$). La covarianza se define como:

$$\text{Cov}(X_i, X_j) = E[(X_i - E[X_i])(X_j - E[X_j])], \quad (1.4)$$

o bien su versión normalizada, la función de correlación cruzada que siempre toma valores comprendidos en el intervalo $[-1, 1]$:

$$C(X_i, X_j) = \frac{\text{Cov}(X_i, X_j)}{\sqrt{\text{Var}(X_i)\text{Var}(X_j)}}, \quad (1.5)$$

donde $\text{Var}(\cdot)$ denota la varianza. Un caso particular de la función de correlación cruzada y que vamos a analizar de manera larga y extendida en el transcurso de este trabajo es la *función de autocorrelación*. La función de autocorrelación estudia la dependencia o independencia lineal de una señal consigo misma tras haberla sometido a un desplazamiento temporal k . La función de autocorrelación de una variable X a distancia k puede expresarse como:

$$C_X(k) = \frac{\text{Cov}(X_i, X_{i+k})}{\sqrt{\text{Var}(X_i)\text{Var}(X_{i+k})}} \quad (1.6)$$

1.1.2. Estacionariedad

La estacionariedad de una serie temporal $\{x_i\}_{i=1, \dots, N}$ es una medida de la regularidad de la serie a lo largo del tiempo. La *estacionariedad estricta* se consigue únicamente

cuando la distribución de probabilidad conjunta (Eq. 1.2) es idéntica a la que se obtiene al realizar cualquier desplazamiento temporal admisible k en la serie temporal para todo subconjunto $\{x_j\}_{j=i, \dots, i+m}$ de tamaño $m \in \mathbb{N}$. Sin embargo, esta definición es demasiado estricta cuando trabajamos con series temporales procedentes de sistemas reales, por lo que de ahora en adelante cuando mencionemos el concepto de estacionariedad entenderemos que los momentos de orden 1 y 2 (media y varianza respectivamente) de la serie temporal permanecen constantes a lo largo del tiempo (estacionariedad débil o en el sentido débil).

1.2. Ruido Fraccionario Gaussiano y Movimiento Fraccionario Browniano

Desde tiempos ancestrales el Río Nilo ha jugado un papel importantísimo para la supervivencia de los pueblos que habitaban las regiones bañadas por sus aguas. El hecho de que el nivel de sus aguas muestre un comportamiento de *largo alcance* ha sido constatado por muchos historiadores a lo largo de los años. Largos periodos de lluvias eran seguidos por también largos períodos de sequías, con la importancia que esto conllevaba para la agricultura y la ganadería. El hidrólogo Harold Edwin Hurst (1951) analizó series temporales en el flujo de agua del Nilo y se percató de que la serie temporal tenía un aspecto muy estacionario (Beran, 1994). Sin embargo, al analizar pequeños intervalos de tiempo la serie presentaba ciclos y tendencias. Este efecto fue denominado como ‘*efecto de Hurst*’ o ‘*efecto Joseph*’ y puede ser explicado en términos de la capacidad ideal de un embalse (Mandelbrot, 1979; Mandelbrot and Pignoni, 1983; Mandelbrot and van Ness, 1968; Mandelbrot and Wallis, 1968a,b). Supongamos que queremos calcular la capacidad ideal de un embalse en el periodo de tiempo comprendido entre t y $t + k$. Con objeto de simplificar, asumimos que el tiempo es discreto y que el embalse no presenta pérdidas debidas a evaporaciones, fugas, etc. Cuando hablamos de capacidad ideal entendemos que el flujo de salida debe ser uniforme y que en el instante temporal $t + k$ el embalse debe contener la misma cantidad de agua que en el instante inicial t . Además, el embalse en ningún momento podría desbordarse. Denotemos por X_i al flujo de entrada en el instante temporal i de forma que $Y_j = \sum_{i=1}^j X_i$ sea el flujo de entrada acumulado entre los instantes temporales i y j . Entonces, la capacidad ideal del embalse es igual a:

$$R(t, k) = \max_{0 \leq i \leq k} [Y_{t+i} - Y_t - \frac{i}{k}(Y_{t+k} - Y_t)] - \min_{0 \leq i \leq k} [Y_{t+i} - Y_t - \frac{i}{k}(Y_{t+k} - Y_t)]. \quad (1.7)$$

Con objeto de estudiar las propiedades que presentan invariancia frente a escala, se normaliza $R(t, k)$ mediante el factor:

$$S(t, k) = \sqrt{k^{-1} \sum_{i=t+1}^{t+k} (X_i - \bar{X}_{t,k})^2}, \quad (1.8)$$

donde $\bar{X}_{t,k} = k^{-1} \sum_{i=t+1}^{t+k} X_i$. El ratio R/S es llamado ‘*rango reajustado reescalado*’ o simplemente estadístico R/S . Hurst representó el estadístico R/S para muchos valores de k , obteniendo un comportamiento en forma de ley de potencias con exponente $H > 1/2$ donde H denota al exponente de Hurst. Este resultado estaba en clara contradicción con los resultados obtenidos mediante procesos de Markov y otros procesos estocásticos usados para el modelado de series temporales en aquel tiempo, donde el estadístico R/S mostraba una dependencia en forma de ley de potencias con $H = 1/2$.

Entre los muchos modelos propuestos para la generación sintética de series temporales que reproducen el *efecto de Hurst* destaca el ruido fraccionario Gaussiano (fGn por sus siglas en inglés) (Mandelbrot and Wallis, 1968b, 1969a,b). Para la definición del fGn es necesaria la introducción del concepto de *auto-similaridad* (ver (Beran, 1994)). Se entiende por proceso *auto-similar* a todo proceso estocástico Y_t con parámetro de *auto-similaridad* H , en el que para todo factor de escala $c > 0$ el proceso reescalado $c^{-H}Y_{ct}$ tiene una distribución de valores idéntica a la del proceso original Y_t . Supongamos un proceso Y_t *auto-similar* con incrementos $X_i = Y_i - Y_{i-1}$ estacionarios y de media nula. La autocovarianza del proceso con desviación estándar σ viene dada por:

$$\text{Cov}(k) = \frac{1}{2}\sigma^2[(k+1)^{2H} - 2k^{2H} + (k-1)^{2H}], \quad (1.9)$$

para $k \geq 0$ y $\text{Cov}(k) = \text{Cov}(-k)$ para $k < 0$. Entonces, teniendo en cuenta que el proceso Y_t tiene media nula, la función de autocorrelación vendrá dada por:

$$C(k) = \frac{1}{2}[(k+1)^{2H} - 2k^{2H} + (k-1)^{2H}], \quad (1.10)$$

para $k \geq 0$ y $C(k) = C(-k)$ para $k < 0$. Desarrollamos en serie de Taylor $C(k)$ para obtener su comportamiento asintótico. En primer lugar consideramos que $C(k) = \frac{1}{2}k^{2H}g(k^{-1})$ y $g(x) = (1+x)^{2H} - 2 + (1-x)^{2H}$. Si $(0 < H \leq 1)$ y $H \neq \frac{1}{2}$ el primer término no nulo del desarrollo en serie de Taylor de la función $g(x)$ en torno al origen es igual a $2H(2H-1)x^2$ por lo que:

$$C(k) \simeq H(2H-1)k^{2H-2}. \quad (1.11)$$

Asumiendo X_t como un proceso Gaussiano, la distribución de valores estará perfectamente definida por la media y la autocovarianza. Además, para cada valor de $H \in (0, 1)$ existe un único proceso Gaussiano X_t , que es el incremento estacionario del proceso *auto-similar* Y_t . El proceso X_t es el fGn mientras que el proceso Y_t es el *movimiento fraccionario Gaussiano* (fBm por sus siglas en inglés) (Mandelbrot and van Ness, 1968). El clásico movimiento Browniano introducido por Robert Brown en 1828 corresponde al proceso Y_t con $H = \frac{1}{2}$, mientras que su incremento, X_t se denomina *ruido blanco* y ha sido de gran importancia a la hora de modelar la componente puramente aleatoria de procesos reales.

Además, considerando la Eq. 1.11 podemos establecer la aproximación:

$$C(k) \simeq \frac{H(2H-1)}{k^{2-2H}} \propto \frac{\text{sign}(1-\gamma)}{k^\gamma}, \quad (1.12)$$

donde γ es el exponente de caída de la función de autocorrelación, dado por $\gamma = 2 - 2H$ con $\gamma \in (0, 2)$. Por tanto, para valores $\gamma < 1$ ($H > 0.5$) las correlaciones son positivas y de largo alcance, ya que el exponente de caída es menor que la unidad, mientras que para valores $\gamma > 1$ ($H < 0.5$) las correlaciones serán negativas y de corto alcance (también llamadas anticorrelaciones) debido a que la función de autocorrelación decae mediante una ley de potencias con exponente de caída mayor que la unidad. Para el caso $\gamma = 0$ ($H = 0.5$) la función de autocorrelación se anula y no existen correlaciones (ruido blanco).

1.3. Técnicas de Análisis de Fluctuación

Como hemos visto en la sección previa, la función de autocorrelación determina únicamente las propiedades de correlación lineal de una serie temporal para diferentes retardos temporales. Sin embargo, en muchas ocasiones la función de autocorrelación es ruidosa y sensible al tamaño del sistema, dificultando por tanto su estudio. Debido a esto, se ha estandarizado el uso de técnicas indirectas para cuantificar el exponente de autocorrelación, siendo el Fluctuation Analysis (FA) y el Detrended Fluctuation Analysis (DFA) las más utilizadas. Estos métodos estudian la fluctuación de la señal a distintas escalas, logrando un resultado mucho más *suaizado* y parecido a una ley de potencias que el logrado mediante la función de autocorrelación, enmascarando por tanto correlaciones que no decaen mediante una ley de potencias.

1.3.1. Fluctuation Analysis

El FA (Bryce and Sprague, 2012; Peng et al., 1992) es una técnica que se utiliza para el estudio del scaling¹ de series estacionarias. En primer lugar, dada una serie estacionaria x_i se calcula su serie acumulada $Y(j) = \sum_{i=1}^j x_i$.

La fluctuación para una determinada ventana de tamaño ℓ viene dada por:

$$F_{\text{FA}}(\ell) = \sqrt{\langle (Y_{i+\ell} - Y_i)^2 \rangle} \quad (1.13)$$

El scaling existirá única y exclusivamente cuando $F_{\text{FA}}(\ell) \sim \ell^\alpha$. El exponente α se calcula mediante un ajuste lineal de $\log(F_{\text{FA}}(\ell))$ vs. $\log(\ell)$, y está relacionado con el exponente de autocorrelación γ mediante (Allegrini et al., 1995):

$$\gamma = 2 - 2\alpha \quad (1.14)$$

Por tanto, teniendo en cuenta la Eq. 1.12, para señales correlacionadas estacionarias, $\alpha \in (0, 1)$ y coincide con el exponente de Hurst (H). Valores $\alpha < 0.5$ indican la presencia de correlaciones negativas o anticorrelaciones, mientras que el valor $\alpha = 0.5$ indica ausencia de correlaciones (ruido blanco). Además, mediante el Teorema de Wiener-Khinchin (Rangarajan and Ding, 2000), se puede probar que cuando la función de autocorrelación de una serie temporal sigue una ley de potencias ($C(\ell) \sim \ell^{-\gamma}$), su espectro de potencias también lo hace ($S(f) \sim f^{-\beta}$). El exponente de caída del espectro de potencias β está relacionado con α mediante la relación:

$$\alpha = \frac{\beta + 1}{2} \quad (1.15)$$

Las figuras (1.1) y (1.2) muestran los exponentes de correlación α y β obtenidos mediante el FA y el espectro de potencias de la señal para un fGn correlacionado y otro anticorrelacionado respectivamente. En ambos casos, los exponentes γ , α y β toman los valores establecidos por las Eqs. 1.14 y 1.15.

1.3.2. Detrended Fluctuation Analysis

El DFA (Peng et al., 1994b) es una versión modificada del FA que es capaz de eliminar los efectos producidos por la no estacionariedad de la serie temporal. De manera análoga al FA, este método ofrece un parámetro de salida α para cuantificar las propiedades de scaling de series con correlaciones de largo alcance. Por tanto, mediante el uso del DFA

¹Se entiende por scaling a la invariancia frente a cambios de escala, asociada a una forma funcional de ley de potencias.

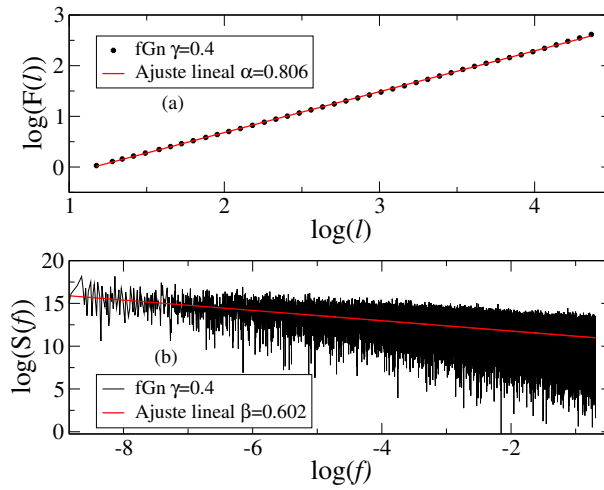


Figura 1.1: a) FA para un ruido fraccionario Gaussiano con exponente de autocorrelación $\gamma = 0.4$ y tamaño de secuencia $N = 2^{18}$. El exponente FA obtenido es $\alpha = 0.806$. b) Espectro de potencias del ruido fraccionario Gaussiano utilizado en (a). El exponente de caída del espectro de potencias es $\beta = 0.602$.

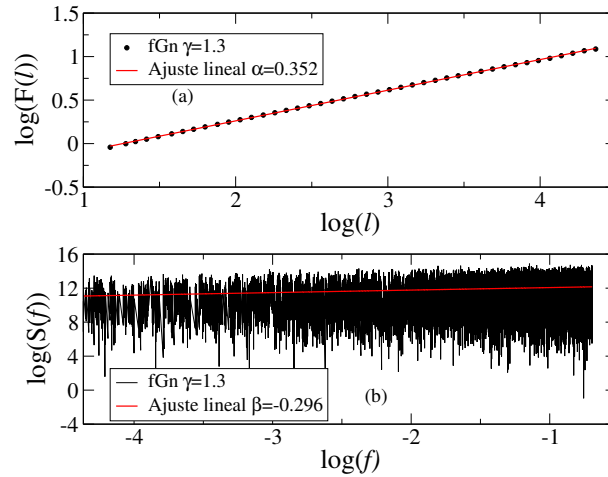


Figura 1.2: a) FA para un ruido fraccionario Gaussiano anticorrelacionado con exponente de autocorrelación $\gamma = 1.3$ y tamaño de secuencia $N = 2^{18}$. El exponente FA obtenido es $\alpha = 0.352$. b) Espectro de potencias del ruido fraccionario Gaussiano utilizado en (a). El exponente de caída del espectro de potencias es $\beta = 0.296$.

es posible cuantificar las propiedades de scaling tanto de fGns ($\alpha < 1$) como de fBms ($\alpha \geq 1$). El algoritmo DFA se puede resumir en los siguientes pasos:

- (i) Partiendo de una serie correlacionada x_i de tamaño N obtiene la señal integrada

$$y(j) \equiv \sum_{i=1}^j [x_i - \mu] \quad (1.16)$$

donde μ es la media de la serie.

- (ii) La serie integrada $y(j)$ se divide en *ventanas* de igual tamaño ℓ .
- (iii) En cada ventana de tamaño ℓ se calcula un ajuste lineal de $y(j)$ que representa la *tendencia lineal* dentro de esa ventana. La coordenada y de la recta de ajuste en cada venta se denota por $y_\ell(j)$.
- (iv) Se elimina la tendencia a la serie integrada $y(j)$ mediante la resta de la tendencia local en cada ventana $y_\ell(j)$ de longitud ℓ .
- (v) Para un determinado tamaño de ventana ℓ , se calcula la raíz cuadrática media $F_{\text{DFA}}(\ell)$ (r.m.s por sus siglas en ingles) de la señal integrada una vez eliminada la tendencia:

$$F_{\text{DFA}}(\ell) = \sqrt{\frac{1}{N} \sum_{j=1}^N [y(j) - y_\ell(j)]^2} \quad (1.17)$$

- (vi) El cálculo efectuado en la Eq. 1.17 se repite para un gran rango de tamaños de ventana ℓ para obtener la relación entre $F(\ell)$ y el tamaño de ventana ℓ .

Para una serie con correlaciones (tanto de largo como de corto alcance) se debe cumplir que la fluctuación escale como una ley de potencias en función del tamaño de ventana ($F_{\text{DFA}}(\ell) \sim \ell^\alpha$). En el caso del DFA, la serie temporal presenta anticorrelaciones si $\alpha > 0.5$, es ruidosa o no correlacionada si $\alpha = 0.5$ y presenta correlaciones de largo alcance si $\alpha \in (0.5, 2)$. En particular, para el caso $\alpha = 1.5$ la serie corresponde al Movimiento Browniano, que como ya se mencionó previamente en la Sec. 1.2 coincide con la integral del ruido blanco.

El DFA ha sido sistemáticamente estudiado con series temporales con diferentes tipos de tendencias (Hu, 2001; Xu, 2005), pérdida de datos (Ma, 2010), diferentes artefactos (Chen, 2002), pre-procesados lineales y no lineales (Chen, 2005) y coarse-graining de la serie temporal (Xu et al., 2010). Aunque el DFA está muy estandarizado, un uso incorrecto del mismo puede llevar a la consecución de resultados espurios. Hay que tener en mente

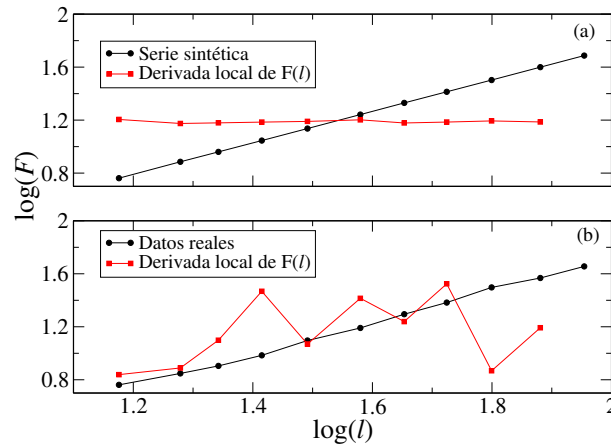


Figura 1.3: Ejemplo de dos series temporales con y sin scaling e idéntico exponente de correlación α obtenido mediante el DFA. (a) Serie sintética. En este caso podemos ver que la derivada local de $F(\ell)$ apenas fluctúa por lo que la señal tiene un buen scaling y el uso del DFA para determinar el exponente de correlación estaría perfectamente justificado. (b) Serie temporal de datos reales, en la que a pesar de mostrar un exponente de correlación α idéntico al obtenido en (a) el scaling no es para nada bueno ya que la derivada local no es constante y presenta grandes fluctuaciones. En este caso, el uso del DFA para la determinación del exponente de autocorrelación podría conducir a resultados espurios.

que el DFA está originalmente diseñado como una técnica de análisis de fluctuaciones, que nos permite calcular de forma indirecta el exponente de autocorrelación de la señal temporal bajo estudio. Dicho esto, el DFA únicamente es aplicable cuando la función de autocorrelación (o la fluctuación en función de la escala) decae mediante una ley de potencias, i.e. hay scaling. Para cerciorarnos sobre este aspecto, una técnica muy útil consiste en analizar la derivada local del *plot* en doble escala logarítmica de la curva $F(\ell)$ vs. ℓ . En la Fig. 1.3 podemos ver un claro ejemplo de dos series (una sintética y otra de datos reales) en las que a pesar de tener el mismo exponente DFA ($\alpha \sim 1.2$) obtenido a través de un ajuste global, presentan propiedades de scaling muy dispares. En la serie sintética el scaling es prácticamente perfecto, ya que la derivada local fluctúa muy poco en torno a un valor constante, mientras que en la serie de datos reales la fluctuación es muy grande aunque el ajuste de $F(\ell)$ sea bueno. Además, en el Cap. 3 veremos que las técnicas de análisis de fluctuación también pueden conducir a resultados espurios en el análisis de correlaciones de series módulo y signo, incluso cuando el scaling es suficientemente bueno.

1.4. Método de Filtrado de Fourier

El método de Filtrado de Fourier (FFM por sus siglas en inglés) (Bernaola-Galván et al., 2012; Makse et al., 1996a) es uno de los algoritmos más utilizados a la hora de

generar de forma sintética fGns y fBms en los que podamos controlar a priori sus propiedades. El método hace uso de la Eq. 1.15 y puede resumirse en los siguientes pasos:

- (i) Generamos un ruido blanco $\eta(i)$, i.e. una serie estacionaria sin correlaciones y con una distribución de valores Gaussiana. Calculamos su transformada de Fourier $\hat{\eta}(f)$.
- (ii) La serie con el exponente de correlación α deseado se obtiene como

$$x(i; \alpha) \equiv \mathcal{F}^{-1} \left[\frac{\hat{\eta}(f)}{f^{\alpha-1/2}} \right], \quad (1.18)$$

donde $\mathcal{F}^{-1}[\cdot]$ denota la transformada inversa de Fourier. Para comprobarlo, simplemente se debe tener en cuenta que la transformada de Fourier de $x(i; \alpha)$ es una ley de potencias de exponente $\alpha - 1/2$ y por tanto, su espectro de potencias sigue una ley de potencias con exponente $2\alpha - 1$ por lo que, de acuerdo con Eq. 1.15 el exponente DFA sería α . Las series generadas mediante el FFM tienen media nula y desviación estándar unidad.

Otros autores resuelven el problema de crear series sintéticas con propiedades controladas a priori a partir de derivaciones e integraciones fraccionales de ruidos blancos (Herrmann, 2014; Hosking, 1981; Wyss, 1991). Denotemos $\lambda = \alpha - 1/2$. De acuerdo con Eq. 1.18 valores positivos de λ pueden ser considerados como integraciones fraccionales de orden λ (no necesariamente entero) de $x(i; \alpha)$, mientras que valores negativos de λ son vistos como derivadas fraccionales de orden $-\lambda$. Por tanto, los ruidos fraccionarios Gaussianos corresponden a valores $\lambda < 0.5$ y los movimientos fraccionarios Brownianos a valores $\lambda \geq 0.5$, incluyendo el ruido $1/f$, también conocido como *ruido rosa* (Lowen and Teich, 2005).

De la Eq. 1.18 podemos concluir que un ruido fraccionario exhibe un comportamiento en forma de ley de potencias inversa en su espectro de potencias, con exponente $\beta = 2\lambda$, y de ahí que a este tipo de ruidos se les conozca también por ruidos $1/f^\beta$. En la Fig. 1.4 aparecen algunos ejemplos de ruidos generados mediante el Método de Filtrado de Fourier.

1.4.1. Generación de ruidos fraccionarios con espectro de potencias limpio

Como ya se ha explicado previamente, el hecho de que una serie temporal tenga una función de autocorrelación en forma de ley de potencias, implica que el espectro de potencias también tendrá una dependencia en forma de ley de potencias con la frecuencia. Además, es posible relacionar el exponente de correlación α del DFA que caracteriza al

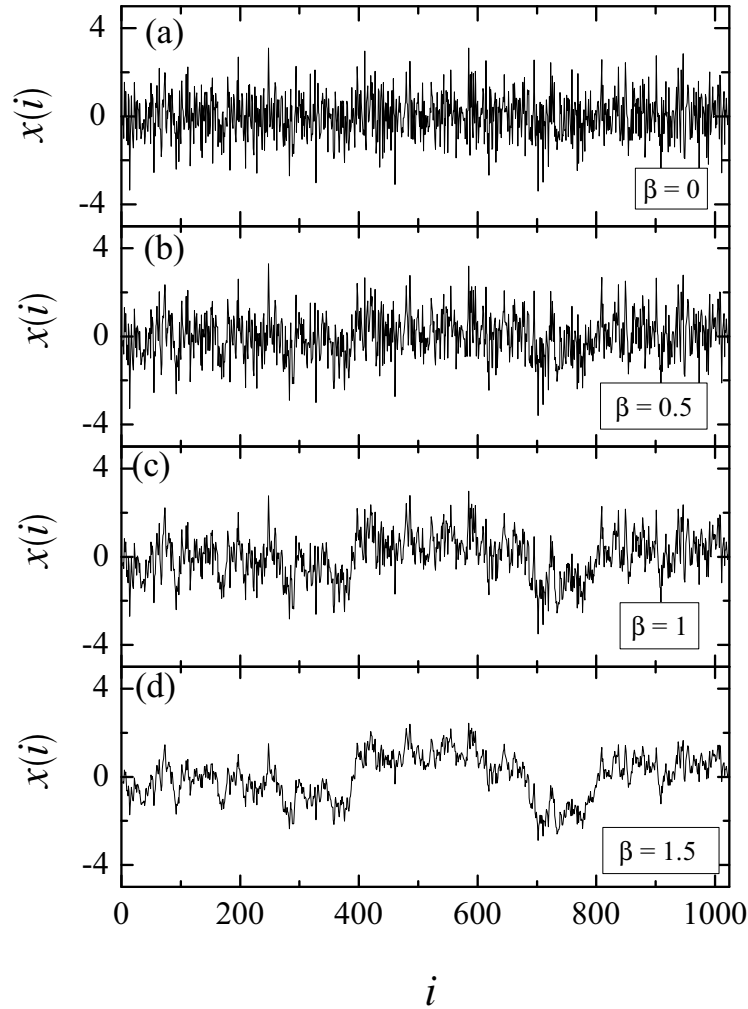


Figura 1.4: Ejemplos de ruidos fraccionarios generados mediante el Fourier Filtering Method. Este método es una aplicación directa de la definición de ruido fraccionario mostrada en la Eq. 1.18: En primer lugar, generamos ruido blanco $\eta(i)$ con distribución de valores Gaussiana y calculamos su transformada de Fourier $\hat{\eta}(f)$. Para obtener la serie con el deseado exponente de correlación β en el espectro de potencias, simplemente dividimos $\hat{\eta}(f)$ en el espacio de Fourier por $f^\lambda = f^{\beta/2}$ y calculamos la transformada inversa de Fourier. Las señales representadas en esta figura tienen un tamaño de $N = 1024$ puntos y todas ellas han sido generadas a partir de la misma serie inicial de ruido blanco Gaussiano. (a) Ruido blanco ($\beta = 0$, $\alpha = 0.5$), (b) ruido fraccionario Gaussiano con ($\beta = 0.5$, $\alpha = 0.75$), (c) ruido $1/f$ ($\beta = 1$, $\alpha = 1$) y (d) movimiento fraccionario Browniano ($\beta = 1.5$, $\alpha = 1.25$) (Figura publicada originalmente en (Bernaola-Galván et al., 2012)).

fGn o fBm en cuestión con el correspondiente exponente del espectro de potencias (Eq. 1.15). Sin embargo, en las Figs. 1.1 y 1.2 vemos que a pesar de que el valor del exponente de ajuste mediante una ley de potencias realizado al espectro de potencias coincide plenamente con el valor que predice la Eq. 1.15, el espectro de potencias presenta un aspecto por lo general bastante ruidoso. Para evitar esto, de Moura (de Moura and Lyra, 1998) introdujo una modificación al método de Filtrado de Fourier que permite la obtención de ruidos fraccionarios Gaussianos y movimientos fraccionarios Brownianos con espectro de potencias totalmente limpio. El algoritmo se resume en los siguientes pasos:

- De forma contraria al método de filtrado de Fourier, donde se partía de un ruido blanco Gaussiano, ahora el punto de partida va a ser la propia transformada de Fourier, definida como:

$$\text{Re}(f) = f^{-\beta/2} \cos(2\pi u) \quad (1.19)$$

$$\text{Im}(f) = f^{-\beta/2} \text{sen}(2\pi u) \quad (1.20)$$

donde $\text{Re}(f)$ e $\text{Im}(f)$ son las partes real e imaginaria de la transformada de Fourier para la frecuencia f y u es un número aleatorio perteneciente a una distribución de valores Gaussiana. Con esta definición de la transformada de Fourier, nos aseguramos que el espectro de potencias siga estrictamente una ley de potencias con exponente de caída β .

- Realizamos la transformada inversa de Fourier para volver al dominio del tiempo, y normalizamos la serie mediante la desviación estándar.

En la Fig. 1.5 se representa el espectro de potencias obtenido mediante esta técnica para ruidos con diferentes exponentes de correlación.

1.5. Multifractal DFA

El DFA estudia el scaling del momento de segundo orden en función del tamaño de ventana ℓ , y consecuentemente, solo considera las correlaciones lineales presentes en la serie temporal. El Multifractal Detrended Fluctuation Analysis (MFDFA por sus siglas en inglés) puede ser entendido como una generalización del DFA que analiza el scaling de todos los posibles momentos de orden q , incluyendo aquellos negativos (Kantelhardt, 2002). La Eq. 1.17 se generaliza de la siguiente forma:

$$F_q(\ell) = \left(\frac{1}{N} \sum_{j=1}^N |y(j) - y_\ell(j)|^q \right)^{\frac{1}{q}} \quad (1.21)$$

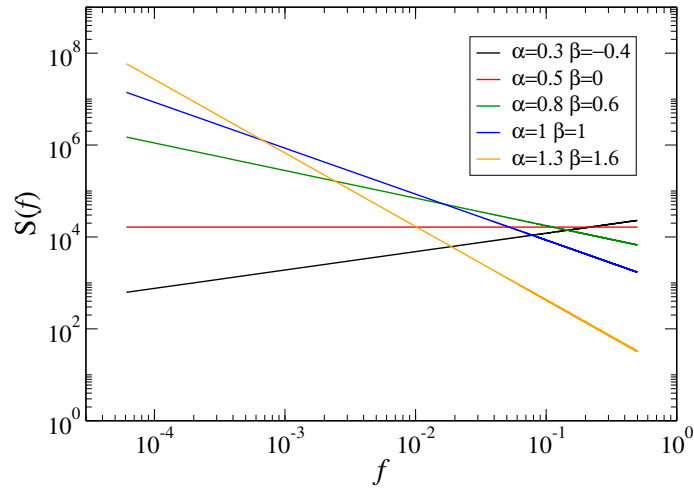


Figura 1.5: Espectros de potencias de series sintéticas de tamaño 2^{14} generadas mediante el método expuesto en (de Moura and Lyra, 1998) con diferentes exponentes de correlación. Como se puede apreciar, los espectros de potencias siguen una rigurosa dependencia en forma de ley de potencias con la frecuencia.

En el caso de series temporales con correlaciones de largo alcance que decaen con un comportamiento de ley de potencias,

$$F_q(\ell) \sim \ell^{h(q)} \quad (1.22)$$

donde $h(q)$ es el exponente de scaling de las fluctuaciones de orden q en función del tamaño de ventana ℓ . Obviamente, el exponente DFA α es un caso particular para $q = 2$, i.e. $\alpha = h(2)$. Para series temporales en las que únicamente hay presentes correlaciones lineales $h(q) = \alpha \quad \forall q$, es decir, existe un único exponente de scaling y la serie se denomina *monofractal*. En cambio, cuando la serie posee correlaciones no lineales cada momento de orden q escala con un exponente $h(q)$ diferente y la serie tiene carácter *multifractal*.

Los exponentes de scaling $h(q)$ se pueden relacionar con los exponentes clásicos multifractales $\tau(q)$ mediante la expresión:

$$\tau(q) = qh(q) - 1 \quad (1.23)$$

Finalmente, calculando la Transformada de Legendre podemos obtener el espectro multifractal (ver (Kantelhardt, 2002)).

$$\zeta = \tau'(q) \quad (1.24)$$

$$f(\zeta) = q\zeta - \tau(q) \quad (1.25)$$

donde $f(\zeta)$ denota la dimensión fractal del subconjunto de la serie temporal caracterizado

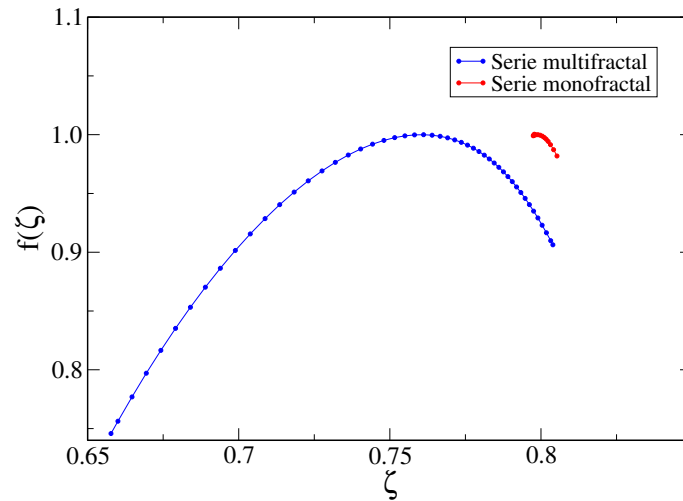


Figura 1.6: Espectros multifractales obtenidos mediante el MFDFA para una serie temporal monofractal (fGn generado mediante el método de Filtrado de Fourier con $\alpha = 0.8$) y otra multifractal generada mediante composición de módulos y signos independientes (ver Cap. 2), ambas de tamaño $N = 2^{18}$. En el caso de la serie monofractal la anchura del espectro multifractal es prácticamente nula, mientras que para la serie multifractal la anchura del espectro es considerable, indicando por tanto la presencia de correlaciones no lineales.

por ζ . Para el caso particular de series monofractales, debido a que todos los momentos de orden q escalan con el mismo exponente, $h(q) = \alpha$, $h'(q) = 0$ y el espectro multifractal tendrá un aspecto de función delta:

$$f(\zeta) = \delta(\zeta - \alpha). \quad (1.26)$$

Por el contrario, para una serie temporal multifractal, $f(\zeta)$ tendrá una anchura $\Delta\zeta$ distinta de cero (ver Fig. 1.6), la cual puede usarse como medida de la no-linealidad presente en la serie temporal.

1.6. Transformada Wavelet

La transformada Wavelet (ver (Arneodo et al., 2002))(WT) es una técnica introducida en los años ochenta por (Goupillaud et al., 1984; Grossmann and Morlet, 1985) que permite la expansión de series temporales en términos de *wavelets* construidas a partir de una *función madre*, ψ , por medio de traslaciones y dilataciones. La transformada wavelet de una función real $f(x)$ se define como:

$$T_\psi[f](x_0, a) = \frac{1}{a} \int_{-\infty}^{+\infty} f(x) \psi\left(\frac{x - x_0}{a}\right) dx \quad (1.27)$$

donde x_0 es el parámetro espacial y a ($a > 0$) es la escala de observación. Generalmente, la función ψ es una función bien localizada tanto en tiempo como en frecuencia, de media nula y ortogonal a polinomios de bajo orden, es decir:

$$\int_{-\infty}^{+\infty} x^m \psi(x) dx = 0, \quad \forall m, \quad 0 \leq m \leq n_\psi, \quad m \in \mathbb{N} \quad (1.28)$$

Desde su introducción ha sido objeto de estudio y aplicada en diferentes campos (Arneodo and Bacry, 1995; Coifman et al., 1992; Erlebacher et al., 1996). La WT es considerada un microscopio matemático ya que sus coeficientes son una expresión del comportamiento de la serie temporal a distintos niveles de resolución (escalas) (Arneodo et al., 1988). A modo de ejemplo, en la Fig. 1.7 podemos ver cómo al aplicar la WT a una serie temporal, esta se suaviza a medida que aumentamos la escala de observación. Además, varios trabajos muestran que toda la información relevante que puede aportar la WT está contenida en su *skeleton* o líneas de módulo máximo (Mallat et al., 1992). Dichas líneas están definidas para cada valor de la escala a como los máximos locales de $|T_\psi[f](x, a)|$, considerados como una función de la variable espacial x . El interés en las líneas de máximo reside en que escalan como:

$$T_\psi[f](x_0, a) \sim a^{h(x_0)} \quad (1.29)$$

siendo $h(x_0)$ el exponente de Hurst local (también conocido como el exponente de Hölder) en el punto espacial x_0 . Esta técnica nos permite evaluar el exponente de Hurst (exponente de correlación) local en cada punto de la serie temporal, y por tanto, mediante una forma totalmente diferente a la descrita en Sec. 1.5 podemos obtener de nuevo información concluyente acerca de las propiedades multifractales del sistema bajo estudio. Además, puede ser de gran utilidad a la hora de evaluar el comportamiento local de ciertas series temporales procedentes de sistemas presentes en la naturaleza, donde en muchos casos las técnicas de análisis de fluctuación (FA y DFA) no pueden ofrecer buenos resultados debido a la ausencia de scaling y al comportamiento altamente no estacionario de la serie temporal.

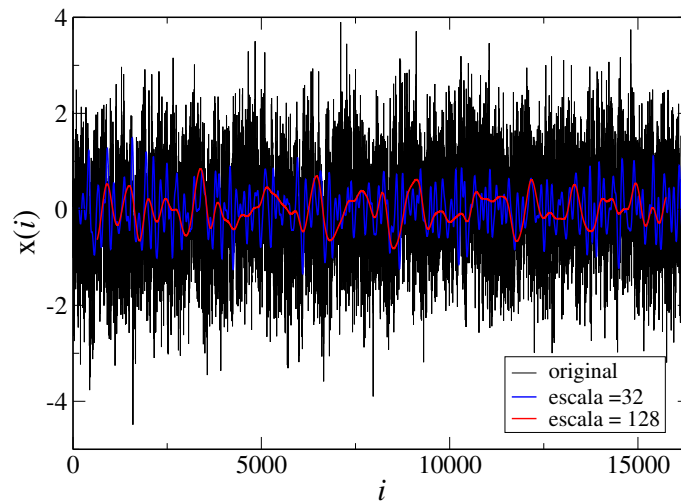


Figura 1.7: Serie sintética de tamaño $N = 2^{14}$ y transformada Wavelet mediante la primera derivada Gaussiana para escalas $a = 32$ y $a = 128$.

1.7. Randomización de fases

La randomización de fases (Ivanov, 1999a; Schreiber and Schmitz, 2000) es una técnica que permite eliminar las no-linealidades presentes en una serie temporal. Dada una serie temporal con no-linealidades, $\eta(i)$, el proceso consiste en:

- (i) Calcular la Transformada de Fourier $\eta(j) = \mathcal{F}(\eta(i))$
- (b) Modificamos $\eta(j)$, conservando las amplitudes y barajando las fases, dando lugar a $\hat{\eta}(j)$. De esta forma nos cercioramos que las correlaciones lineales no se verán afectadas, ya que el espectro de potencias únicamente depende de la amplitud. En cambio, las no-linealidades presentes en la serie serán eliminadas.
- (c) Calculamos la Transformada inversa de Fourier para volver al dominio del tiempo: $\hat{\eta}(i) = \mathcal{F}^{-1}(\hat{\eta}(j))$

En la Fig. 1.8 se muestra la pérdida de no-linealidad en una serie temporal tras realizarse un barajado de fases. El espectro multifractal reduce muy considerablemente su anchura una vez realizado el barajado de fases.

1.8. Transformación a distribución Gaussiana

En algunos casos, dada una serie temporal con una distribución de valores no Gaussiana, $\{x_i\}$, es necesario convertirla a Gaussiana, $\{x'_i\}$. Para ello, se puede hacer uso de

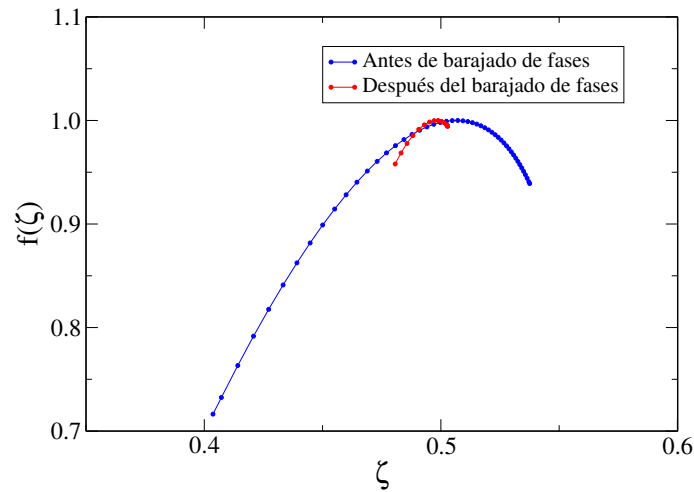


Figura 1.8: Espectros multifractales obtenidos mediante el MFDFA para una serie temporal multifractal antes y después de realizar el barajado de fases. El tamaño de la serie es $N = 2^{18}$.

la siguiente transformación:

$$x' = \Phi^{-1} [F(x)] , \quad (1.30)$$

donde $F(\cdot)$ es la distribución acumulada de la serie de partida $\{x_i\}$ y $\Phi(\cdot)$ es la distribución acumulada de la distribución normal estándar $\mathcal{N}(0, 1)$. Además, hemos comprobado que esta transformación apenas modifica las correlaciones lineales de la serie de partida. La Fig. 1.9 muestra los pasos a seguir para realizar esta transformación, partiendo de una serie cuya distribución de valores no es Gaussiana (en este caso se trata de una exponencial, pero es extrapolable a cualquier otra distribución). Además, las correlaciones lineales de la serie de partida no se ven excesivamente afectadas tras esta transformación. La Fig. 1.10 muestra el DFA de dos series sintéticas: (i) serie con exponente de correlación $\alpha = 0.8$ y distribución de valores exponencial, (ii) la misma serie usada en (i) una vez realizada la transformación, y por tanto, con distribución de valores Gaussiana

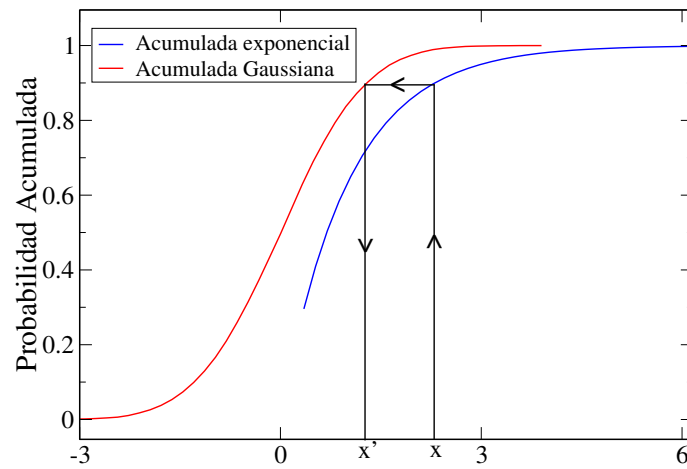


Figura 1.9: Transformación a distribución Gaussiana de una serie cuya distribución de valores original no es Gaussiana. En primer lugar, se calcula la distribución acumulada original de la serie (línea azul) y para cada valor de la serie original $\{x_i\}$ se calcula su correspondiente valor en la distribución acumulada, se mapea ese valor a la distribución acumulada Gaussiana (línea roja) y finalmente se obtiene el valor correspondiente $\{x'_i\}$ que sigue una distribución de valores Gaussiana.

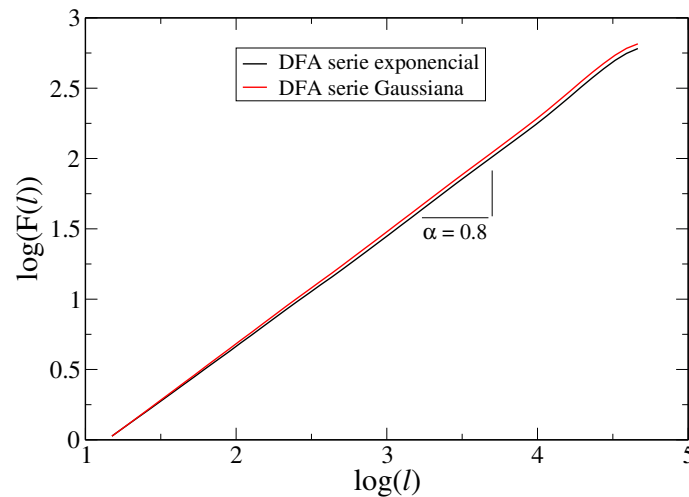


Figura 1.10: DFAs antes y después de la transformación a distribución Gaussiana de una serie con correlaciones de largo alcance $\alpha = 0.8$ cuya distribución inicial de valores es exponencial. Como se puede apreciar, el hecho de realizar la transformación prácticamente no afecta al exponente de correlación, que sigue siendo $\alpha \simeq 0.8$.

Capítulo 2

Módulo y signo de series con correlaciones de largo alcance

Una amplia variedad de fenómenos pertenecientes a diferentes campos, entre los que podemos destacar Fisiología, Geología, Economía, Meteorología, etc, presentan una dinámica compleja con señales de salida que aunque en apariencia parecen ruidosas y erráticas, poseen correlaciones de largo alcance y estructura invariante frente a cambios de escala. Además, la presencia de dichas correlaciones ha sido asociada a propiedades relevantes del sistema bajo estudio. Por ejemplo, las correlaciones en la serie de latidos cardíacos cambian drásticamente entre sujetos sanos y enfermos (Ivanov, 1999a) o entre diferentes estados fisiológicos (Ivanov, 1999b).

En muchos casos, dada una serie temporal x_i , $i = 1, 2, \dots, N$, los incrementos $\Delta x_i = x_{i+1} - x_i$ son más relevantes que la propia serie porque las propiedades de los incrementos ofrecen información importante sobre la dinámica subyacente del sistema y pueden ayudar a desarrollar modelos útiles que lo describan.

Para sistemas no lineales es importante ir más allá del estudio de las correlaciones lineales, ya que estas no tienen en cuenta todas las propiedades dinámicas de este tipo de sistemas — e.g series de incrementos con las mismas correlaciones lineales pueden corresponder a sistemas con comportamientos no lineales y multifractales totalmente diferentes (Ashkenazy, 2001). Una simple estrategia para romper esta degeneración consiste en estudiar de forma separada las correlaciones en el módulo (valor absoluto de la serie) y signo (+1 si el valor es positivo y -1 en el caso de que sea negativo) de la serie de incrementos. Las correlaciones en el módulo (también conocido como volatilidad) han sido asociadas a la presencia de correlaciones no lineales y estructura multifractal (Ashkenazy, 2001, 2003; Kalisky et al., 2005), mientras que las correlaciones del signo están únicamente determinadas por las correlaciones lineales de la serie de procedencia del signo (Ashkenazy, 2001,

2003), y han sido estudiadas en el contexto de tiempos de paso en procesos correlacionados con invariancia de escala (Carretero-Campos et al., 2012).

Además, desde un punto de vista intuitivo, las series módulo y signo contienen información complementaria sobre la serie original: el módulo mide cómo de grandes son los cambios y el signo indica la dirección de estos. A modo de ejemplo, en la dinámica del corazón (Ashkenazy, 2001), que se concibe como el resultado de dos fuerzas que compiten (ramas simpática y parasimpática del Sistema Nervioso Autónomo), esto lleva a una compleja variabilidad con propiedades de invariancia frente a escala. Se podría decir que la primera es responsable de lentos incrementos en el ritmo cardíaco (pequeños en módulo y positivos en signo), mientras que la segunda está normalmente asociada con rápidas bajadas del ritmo cardíaco (grandes en módulo y negativas en signo). Otros ejemplos de la utilidad del análisis del módulo y el signo se encuentran en series temporales obtenidas en sistemas de Dinámica de Fluidos (Zhu, 2012), Geología (Makse et al., 1996b,c), Geofísica (Bartos and Jánosi, 2006; Li et al., 2014) y Economía (Liu, 1999).

A pesar de la importancia de las series módulo y signo que acabamos de mencionar, todavía existen cuestiones abiertas: por ejemplo, dada una serie con conocidas correlaciones de largo alcance es clave saber si hay también correlaciones presentes en la series módulo y signo. Esto es lo que de ahora en adelante llamaremos *descomposición*. En principio, un estudio sistemático de este problema es una tarea ardua, ya que las series originales pueden ser de naturalezas muy variopintas, como hemos mencionado previamente. En lugar de estudiar estas series, vamos a estudiar la descomposición de series artificiales que son comúnmente usadas para modelar el comportamiento de series temporales con correlaciones de largo alcance. En particular, vamos a considerar fGns y fBms para modelar series temporales con correlaciones de largo alcance estacionarias y no estacionarias respectivamente.

Una segunda cuestión (todavía abierta en muchos casos) es cómo el módulo y el signo se acoplan para formar la señal producto¹: por ejemplo, en el corazón humano, al analizar los incrementos de la serie RR^2 , no está aún clara la relación entre el módulo (cómo de grande es el cambio) y el signo (dirección del cambio). Obviamente, un estudio sistemático del acoplamiento entre módulo y signo sería de gran interés para mejorar la comprensión de la relación entre ellos y el comportamiento de los mecanismos de control subyacentes. Específicamente, estamos interesados en cómo las correlaciones de la señal producto están controladas por las correlaciones del módulo y del signo, así como el acoplamiento entre ellos. Sin embargo, los mecanismos de acoplamiento del módulo y del signo serán diferentes

¹Al descomponer una señal en módulo y signo, el producto de ambos garantiza la obtención de la señal original o señal de procedencia

²La serie RR es una serie temporal cuyos elementos son los periodos entre latidos consecutivos

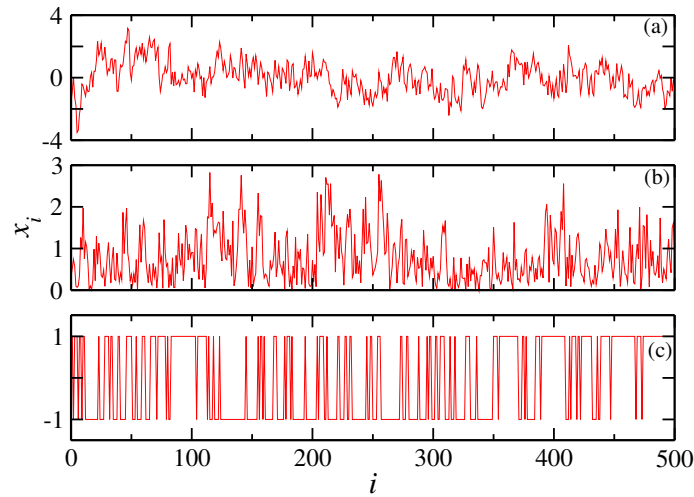


Figura 2.1: (a) Ejemplo de una serie con correlaciones de largo alcance obtenida a través del FFM con $\alpha_{\text{in}} = 1$. (b) Serie módulo y (c) serie signo.

en series temporales de distinta naturaleza, ya que la dinámica subyacente será distinta también y esta variedad de mecanismos potenciales de acoplamiento dificulta un análisis sistemático. En cambio, podemos enfocar este problema desde un punto de vista diferente: podemos estudiar sistemáticamente las correlaciones de series temporales con módulo y signo *desacoplados*. Este enfoque es lo que llamaremos de ahora en adelante *composición*. La *composición* puede ser útil para comprender el comportamiento de sistemas complejos caracterizados por el acoplamiento de dos mecanismos diferentes, cada uno controlando la dinámica del módulo y el signo respectivamente. Además, los resultados de la *composición* pueden ser usados para detectar la existencia de mecanismos de acople al analizar sistemas complejos reales y proporcionan un modelo de series Gaussianas no-lineales.

2.1. Descomposición

Nuestro objetivo es cuantificar las correlaciones lineales en las series módulo y signo (mediante el DFA) obtenidas a través de la descomposición de señales correlacionadas, con una correlación controlada mediante un exponente de correlación de entrada. Para estudiar sistemáticamente la correlación de las señales módulo y signo generamos series artificiales con una longitud $2^{20} \simeq 10^6$ a través del FFM con exponentes de correlación de entrada en el intervalo $(0, 2)$. En la Fig 2.1 se muestra un ejemplo de una serie correlacionada obtenida para $\alpha_{\text{in}} = 1.0$ así como su serie módulo y signo.

El procedimiento que seguimos puede ser resumido de la siguiente forma: primero, para cada serie temporal obtenemos su correspondiente serie módulo y signo, calculamos los exponentes (α_{mag} y α_{sign} respectivamente) y los promediamos una vez realizado este

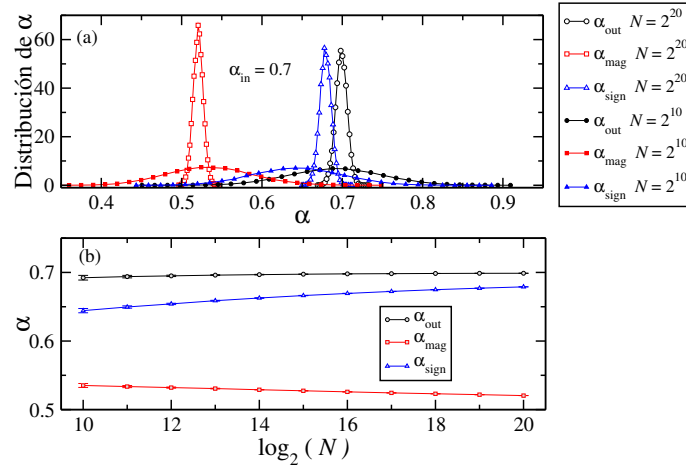


Figura 2.2: *Efectos de tamaño finito*. (a) Distribuciones de α_{out} , α_{mag} y α_{sign} para dos conjuntos de 65000 series generadas a través del Fourier Filtering Method con $\alpha_{in} = 0.7$. Símbolos abiertos: tamaño de la secuencia $N = 2^{20} \simeq 10^6$, símbolos opacos: tamaño de la secuencia $N = 2^{10} \simeq 10^3$. Debido a fluctuaciones estadísticas, los valores obtenidos para el exponente de correlación no son exactamente iguales a α_{in} . De hecho, están distribuidos de acuerdo a una distribución normal con una varianza que decrece conforme N aumenta. Un comportamiento similar es observado para α_{mag} y α_{sign} . (b) Valores medios y desviaciones estándar (barras de error) de las distribuciones de α_{out} , α_{mag} y α_{sign} para experimentos similares a los mostrados en (a) para series temporales con tamaño variando entre $N = 2^{10}$ y $N = 2^{20}$. En los tres casos los valores medios parecen aproximarse a un valor asintótico cuando N aumenta. En particular, α_{out} y α_{sign} tienden al mismo valor α_{in} , siendo la convergencia más lenta para α_{sign} . El hecho de que α_{out} y α_{sign} tengan el mismo límite asintótico se da solo en la región $0.5 \leq \alpha_{in} < 1$, mientras que en la región $\alpha_{in} > 1$, α_{mag} tiende asintóticamente al mismo límite así como α_{out} (ver sección 2.1).

proceso 200 veces para cada exponente de entrada α_{in} . También calculamos el exponente de correlación de la serie temporal generada (α_{out}) que puede ser ligeramente diferente a α_{in} debido a efectos de tamaño finito (ver Fig. 2.2).

En la figura 2.3 representamos los resultados. Podemos claramente distinguir tres regiones:

- (i) $\alpha_{in} \leq 0.5$ A pesar de las anticorrelaciones presentes en la serie temporal, las series módulo y signo son esencialmente ruidosas (sin correlaciones). En todos los casos la serie módulo muestra un ajuste perfecto a una curva en forma de ley de potencias con exponente $\alpha_{mag} = 0.5$ para todas las escalas consideradas. Estas series son prácticamente indistinguibles de las series constituidas por variables aleatorias independientes e idénticamente distribuidas (i.i.d). Por otro lado, el signo (especialmente para $\alpha_{in} > 0.2$) muestra valores de $\alpha_{sign} \lesssim 0.5$, implicando la presencia de anticorrelaciones.

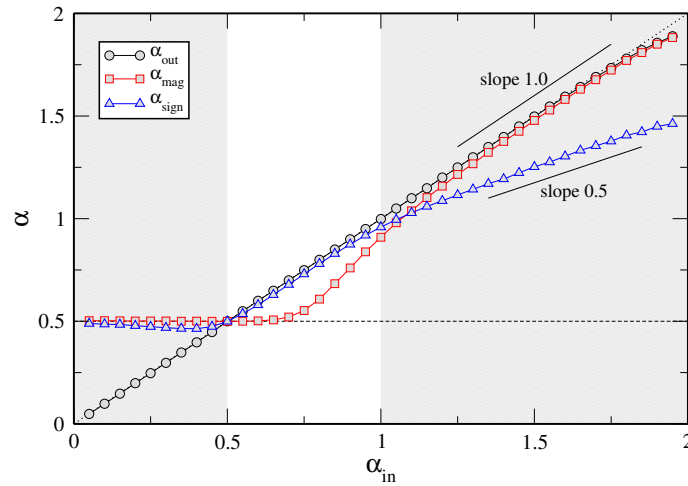


Figura 2.3: Promedio de los exponentes de correlación para la serie compuesta (α_{out}), módulo (α_{mag}) y signo (α_{sign}) en función de α_{in} . Para cada valor de α_{in} generamos 200 series de longitud $N = 2^{20}$ para obtener los valores medios.

Encontramos que el signo únicamente presenta anticorrelaciones a corta escala, mientras que para escalas intermedias o largas muestra claramente un comportamiento ruidoso o no correlacionado. Este efecto se muestra en la Fig. 2.4, donde representamos el comportamiento típico de $F_{\text{sign}}(\ell)$ para el signo en la región $\alpha_{\text{in}} \leq 0.5$. A cortas escalas ℓ , la fluctuación $F_{\text{sign}}(\ell)$ escala con un exponente $\alpha_1 = 0.36$ y, después de un régimen de transición, el resto de la curva muestra un exponente de scaling $\alpha_2 = 0.50$ correspondiente a un comportamiento ruidoso o no correlacionado. Por esta razón, el exponente global α_{sign} , obtenido mediante un ajuste en todo el rango de escalas ℓ (Fig. 2.3), está afectado por estos primeros valores de la curva $F_{\text{sign}}(\ell)$, llevándonos por tanto, a obtener valores de $\alpha_{\text{sign}} \lesssim 0.5$.

En resumen, para escalas suficientemente largas, las series módulo y signo no poseen correlaciones (de nuevo hacemos hincapié en que estos resultados son obtenidos a través del DFA). Teniendo esto en mente, las anticorrelaciones en la serie (presentes en todo el rango de escalas), según el DFA, deben ser fruto del acoplamiento resultante entre el módulo y el signo, porque ninguno de ellos posee anticorrelaciones por sí mismo. Para comprobar esto último, llevamos a cabo el siguiente experimento: Generamos una señal con $\alpha_{\text{in}} = 0.3$, la descomponemos y obtenemos su módulo y su signo, barajamos el signo (destruyendo el posible acoplamiento entre el módulo y el signo) y finalmente multiplicamos el signo aleatorizado por el módulo para obtener una serie subrogada con módulo y signo desacoplados. También repetimos el mismo experimento barajando el módulo en lugar del signo. Los resultados mostrados en la Fig. 2.5 confirman nuestra suposición inicial: las dos series subrogadas pierden

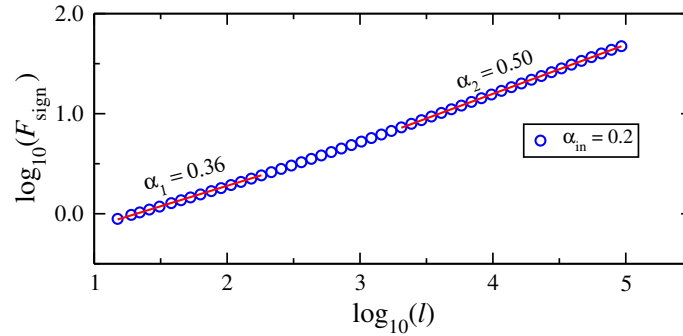


Figura 2.4: $F_{\text{sign}}(\ell)$ vs. ℓ para la serie signo en la región $\alpha_{\text{in}} \leq 0.5$. Promediamos $F_{\text{sign}}(\ell)$ sobre un conjunto de 200 series signo obtenidas a través de series anticorrelacionadas de tamaño $N = 2^{20}$ y $\alpha_{\text{in}} = 0.2$. El exponente de scaling global, $\alpha_{\text{sign}} = 0.46$, indica la presencia de anticorrelaciones, pero mediante una inspección directa de $F_{\text{sign}}(\ell)$ se revela la existencia de un crossover en torno a $\ell_c = 190$. Por debajo de ℓ_c la serie signo muestra principalmente un comportamiento anticorrelacionado ($\alpha_1 = 0.36$), pero tal comportamiento desaparece para $\ell > \ell_c$ ($\alpha_2 = 0.50$).

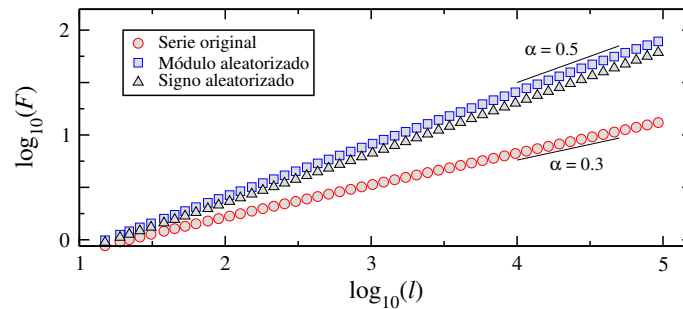


Figura 2.5: $F(\ell)$ vs. ℓ para series anticorrelacionadas ($\alpha_{\text{in}} = 0.3$) y series surrogadas obtenidas mediante la aleatorización del módulo o del signo: Generamos una serie signo con $\alpha_{\text{in}} = 0.3$, la descomponemos y obtenemos sus correspondientes series módulo y signo, aleatorizamos el signo o el módulo y obtenemos dos series surrogadas, una multiplicando el signo aleatorizado por el módulo original (\triangle) y la otra multiplicando el módulo aleatorizado por el signo original (\square). Ambas curvas han sido obtenidas mediante señales con tamaño $N = 2^{20}$ y promediando sobre 200 experimentos.

sus anticorrelaciones ya que $F(\ell)$ escala mediante una ley de potencias $\ell^{0.5}$. Hay que tener en cuenta que en el segundo experimento, donde hemos aleatorizado el módulo, la serie subrogada obtenida aún preserva ciertas anticorrelaciones a corta escala, procedentes de la serie signo original. Sin embargo, para escalas ℓ suficientemente grandes, se recupera el comportamiento aleatorio y las fluctuaciones escalan con $\alpha = 0.5$.

Una importante conclusión que podemos extraer es el hecho de que no podemos obtener secuencias binarias con anticorrelaciones de largo alcance a partir de la serie signo de una serie anticorrelacionada. Esta limitación ha sido también encontrada en otros métodos descritos en la bibliografía para la generación de secuencias binarias con correlaciones de largo alcance (Izrailev et al., 2007; Keshet, 2004; Usatenko et al., 2014).

- (ii) $0.5 < \alpha_{\text{in}} < 1$. La serie signo presenta correlaciones en todo este intervalo, mientras que el módulo únicamente está correlacionado para valores $\alpha_{\text{in}} \geq 0.75$. En cambio, las correlaciones en la señal original están controladas por las del signo, sin importar si el módulo está correlacionado o no. Además, teniendo en cuenta (Apostolov et al., 2008) donde se muestra una relación analítica entre $C(\ell)$ y $C_{\text{sign}}(\ell)$:

$$C(\ell) = \sin \left[\frac{\pi}{2} C_{\text{sign}}(\ell) \right] \quad (2.1)$$

válido para $\gamma < 1$ y $C(\ell) > 0$, i.e. $0.5 < \alpha < 1$. Teniendo en cuenta que las correlaciones tomarán un valor mucho más pequeño que la unidad para el exponente de correlación cuando se consideren escalas suficientemente grandes, el seno en Eq. 2.1 puede ser aproximado por su argumento y, asumiendo una dependencia en forma de ley de potencias para la función de autocorrelación, se tiene que:

$$C(\ell) \approx \ell^{-\gamma} \quad (2.2)$$

$$C_{\text{sign}}(\ell) \approx \ell^{-\gamma_{\text{sign}}}, \quad (2.3)$$

y sustituyendo en Eq. 2.1:

$$\ell^{-\gamma} = \sin \left[\frac{\pi}{2} \ell^{-\gamma_{\text{sign}}} \right]. \quad (2.4)$$

Considerando el primer término del desarrollo en serie de Taylor:

$$\ell^{-\gamma} \approx \frac{\pi}{2} \ell^{-\gamma_{\text{sign}}} \Rightarrow \ell^{\gamma_{\text{sign}} - \gamma} \approx \frac{\pi}{2}. \quad (2.5)$$

Operando:

$$(\gamma_{\text{sign}} - \gamma) \log(\ell) \approx \log\left(\frac{\pi}{2}\right) \Rightarrow \gamma_{\text{sign}} - \gamma \approx \log\left(\frac{\pi}{2}\right) / \log(\ell). \quad (2.6)$$

Por otro lado, usando la Eq. 1.14 tendremos:

$$2(1 - \alpha_{\text{sign}}) - 2(1 - \alpha_{\text{out}}) \approx \log\left(\frac{\pi}{2}\right) / \log(\ell), \quad (2.7)$$

y finalmente:

$$\alpha_{\text{sign}} \simeq \alpha_{\text{out}} - \frac{\log(\pi/2)}{2 \log \ell}, \quad (2.8)$$

donde se relaciona el valor de α_{sign} y α_{out} en función de la escala ℓ .

Se debe tener en cuenta que, de acuerdo con la Eq. 2.8, en la Fig. 2.3 α_{sign} toma siempre valores ligeramente más pequeños que α_{out} . De hecho, $\alpha_{\text{sign}} \rightarrow \alpha_{\text{out}}$ solo de manera asintótica (ver Fig. 2.2.b). Este comportamiento ya ha sido analizado previamente en (Carretero-Campos et al., 2012) estudiando la serie signo en el contexto de la distribución de tiempos de paso en series correlacionadas.

Dentro de esta región, a partir del uso de la serie signo se puede obtener un método para obtener series correlacionadas binarias con un exponente de correlación α_{sign} que es virtualmente igual que el exponente de la serie original α_{in} . Por ejemplo, esté método es útil para estudiar secuencias de ADN que han sido frecuentemente modeladas como secuencias binarias correlacionadas con exponentes de correlación comprendidos en el rango $0.5 < \alpha < 1$ (Bernaola-Galván et al., 2002).

Aquí, de manera contraria a lo que observamos en la región estudiada anteriormente, el acoplamiento entre la series módulo y signo parece no jugar un papel relevante. Como demostraremos en la sección 2.5, incluso bajo la asunción de la independencia entre el módulo y el signo, las correlaciones de la señal producto están controladas por las correlaciones del signo, siempre que se cumpla que $\alpha_{\text{mag}}, \alpha_{\text{sign}} \in (0.5, 1)$.

- (iii) $1 \leq \alpha_{\text{in}} \leq 2$. Ambos α_{mag} y α_{sign} aumentan su valor cuando lo hace α_{in} . Ahora, α_{mag} es el único que tiende asintóticamente a α_{out} , i.e. en esta región las correlaciones de la serie producto están controladas por el módulo. Por otro lado, α_{sign} crece como función de α_{in} con una pendiente $1/2$, por lo que, en esta región:

$$\alpha_{\text{sign}} = \frac{1}{2}(1 + \alpha_{\text{in}}) \quad (2.9)$$

Este comportamiento puede ser explicado de manera analítica usando las propiedades de la distribución de tiempos de paso para series lineales correlacionadas

(Carretero-Campos et al., 2012). También es fácil probar que α_{sign} no puede ser mayor que $3/2$ (Ver Apénd. A para la demostración de ambas propiedades).

En resumen, el exponente de correlación del signo medido mediante el DFA, en función de α_{in} muestra el siguiente comportamiento asintótico:

$$\alpha_{\text{sign}} = \begin{cases} \frac{1}{2} & \alpha_{\text{in}} < \frac{1}{2} \\ \alpha_{\text{in}} & \frac{1}{2} \leq \alpha_{\text{in}} < 1 \\ \frac{1}{2}(1 + \alpha_{\text{in}}) & 1 \leq \alpha_{\text{in}} < 2 \\ \frac{3}{2} & 2 \leq \alpha_{\text{in}} \end{cases} \quad (2.10)$$

En cuanto al exponente de correlación del módulo α_{mag} (medido con el DFA), el comportamiento asintótico consiste en una zona no correlacionada para $\alpha_{\text{in}} < \frac{3}{4}$, un régimen de transición para $\frac{3}{4} < \alpha_{\text{in}} < \frac{5}{4}$ y una región donde $\alpha_{\text{mag}} \simeq \alpha_{\text{in}}$ para $\alpha_{\text{in}} > \frac{5}{4}$. Las correlaciones observadas en la serie módulo coinciden con las encontradas para la serie $(\Delta x_i)^2$ en (Kalisky et al., 2005).

2.2. Composición mediante series módulo y signo

Como ya se dijo en la introducción del presente capítulo, estamos interesados en la composición de series módulo y signo independientes³. Nuestro interés es doble: Por un lado, estudiamos el comportamiento de series temporales cuyo módulo y signo están controlados por mecanismos diferentes. Uno de ellos controla el módulo de los incrementos, mientras que el otro controla el signo (dirección de los incrementos). Esto puede ser considerado como la estrategia más sencilla para el modelado de series reales. Por otro lado, mediante la comprensión de las propiedades de correlación de series compuestas de dichas características, podemos elucidar si módulo y signo de series reales están acoplados o no.

El procedimiento para generar una serie compuesta con módulo y signo independiente se detalla a continuación. Para obtener módulo y signo independientes, usamos el FFM para generar dos series correlacionadas independientes con exponentes de correlación de entrada α_{in1} y α_{in2} , $x(i; \alpha_{\text{in1}})$ y $x(i; \alpha_{\text{in2}})$ respectivamente. Entonces, la serie módulo se obtiene como:

$$x_{\text{mag}}(i) = |x(i; \alpha_{\text{in1}})| \quad (2.11)$$

cuyo exponente de correlación, α_{mag} depende de α_{in1} (Fig. 2.3). De la misma forma, la

³La composición de series módulo y signo independientes se realiza mediante el producto de un módulo por un signo procediendo cada uno de ellos de una serie original distinta.

serie signo se obtiene como:

$$x_{\text{sign}}(i) = \text{sgn}[x(i; \alpha_{\text{in}2})] \quad (2.12)$$

cuyo exponente de correlación, α_{sign} depende de $\alpha_{\text{in}2}$ (Fig. 2.3). Finalmente, la serie producto viene dada por:

$$x_{\text{comp}}(i) = x_{\text{mag}}(i) \cdot x_{\text{sign}}(i) \quad (2.13)$$

Aquí estudiamos sistemáticamente las correlaciones de la serie producto por medio del DFA con $\alpha_{\text{in}1}, \alpha_{\text{in}2}$ en el rango $[0.5, 2]$ ($\alpha_{\text{mag}} \in [0.5, 2]$ y $\alpha_{\text{sign}} \in [0.5, 1.5]$, Fig. 2.3). La región $\alpha_{\text{in}1}, \alpha_{\text{in}2} < 0.5$ no se explora porque, como hemos demostrado previamente en la Sec. 2.1, para estos valores tanto la serie módulo como la serie signo no poseen correlaciones. Dependiendo de los valores de α_{mag} y α_{sign} hemos observado tres comportamientos diferentes:

2.2.1. Caso $\alpha_{\text{sign}} < 1$

Aquí, independientemente del exponente α_{mag} , las correlaciones en la serie producto⁴ están controladas por las correlaciones del signo. Dada una serie temporal obtenida mediante el producto de de dos series independientes módulo y signo, mostramos en el Apéndice B que su función de autocorrelación puede ser escrita como:

$$C(\ell) = C_{\text{sign}}(\ell) \frac{(\pi - 2)C_{\text{mag}}(\ell) + 2}{\pi} \quad (2.14)$$

donde $C(\ell)$, $C_{\text{mag}}(\ell)$ y $C_{\text{sign}}(\ell)$ son las funciones de autocorrelación de la serie compuesta a distancia ℓ , su módulo y su signo respectivamente. Dependiendo de α_{mag} distinguimos dos regímenes:

- (i) $\alpha_{\text{mag}} < 1$. Si la serie temporal presenta correlaciones que decaen en forma de ley de potencias:

$$C_{\text{mag}}(\ell) \sim \ell^{-\gamma_{\text{mag}}} \quad \text{y} \quad C_{\text{sign}}(\ell) \sim \ell^{-\gamma_{\text{sign}}} \quad (2.15)$$

donde $\gamma_{\text{mag}} = 2\alpha_{\text{mag}} - 2$ y $\gamma_{\text{sign}} = 2\alpha_{\text{sign}} - 2$ de acuerdo con (1.14). Usando Eq. 2.15 en Eq. 2.14 obtenemos para la función de autocorrelación de la señal compuesta:

$$C(\ell) \sim \frac{\pi - 2}{\pi} \ell^{-(\gamma_{\text{mag}} + \gamma_{\text{sign}})} + \frac{2}{\pi} \ell^{-\gamma_{\text{sign}}} \quad (2.16)$$

Como estamos considerando $\alpha_{\text{mag}}, \alpha_{\text{sign}} \in [0.5, 1)$, es evidente que $\gamma_{\text{sign}} < \gamma_{\text{mag}} + \gamma_{\text{sign}}$ y entonces, el segundo término será el dominante para escalas ℓ suficientemente

⁴También nos podemos referir a ella como serie compuesta a lo largo del texto

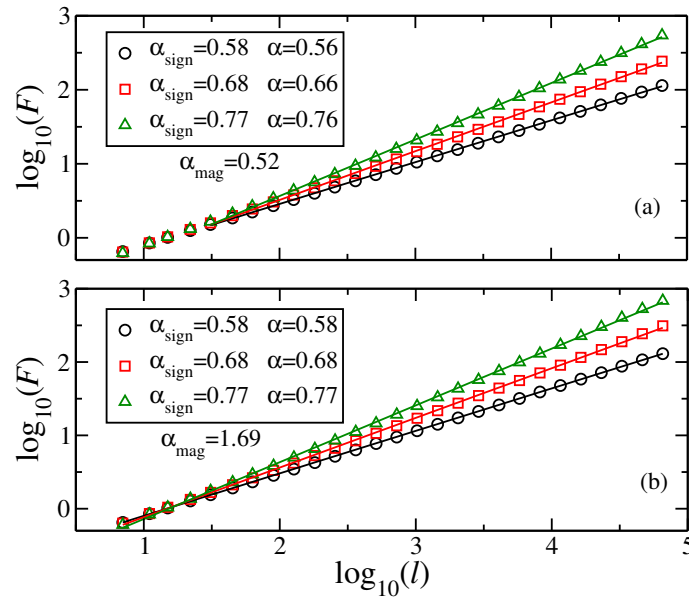


Figura 2.6: Ejemplo de serie compuesta generada mediante la multiplicación de módulo y signo obtenidos a través de series originales independientes para $\alpha_{\text{sign}} < 1$. (a) $\alpha_{\text{mag}} < 1$. Para ℓ suficientemente grande ($\ell > 30$) $F(\ell)$ escala con $\alpha \simeq \alpha_{\text{sign}}$ de acuerdo con Eq. 2.17. (b) $\alpha_{\text{mag}} > 1$. $F(\ell)$ escala con $\alpha = \alpha_{\text{sign}}$ en el rango completo. De forma contraria a la Eq. 2.17, la Eq. 2.18 no es una aproximación para grandes ℓ . El tamaño de la serie es 2^{20} y los resultados se promedian realizando 200 experimentos.

grandes:

$$C(\ell) \sim C_{\text{sign}}(\ell) \quad (2.17)$$

(ii) $\alpha_{\text{mag}} \geq 1$. Ahora $C_{\text{mag}}(\ell) = \text{cte}$ y se obtiene de forma directa a partir de la Eq. 2.14:

$$C(\ell) \propto C_{\text{sign}}(\ell) \quad (2.18)$$

Se debe tener en cuenta que mientras que la Eq. 2.17 es una aproximación válida solo para ℓ suficientemente grande, la Eq. 2.18 mantiene su validez en el rango completo. En la Fig. 2.6 mostramos un ejemplo de dichas situaciones. Para un valor fijado α_{mag} obtenemos series compuestas con diferentes valores de α_{sign} y, en todos los casos, el exponente de correlación resultante toma un valor prácticamente igual al de α_{sign} de acuerdo con las Eqs. 2.17 y 2.18. Teniendo esto último en cuenta, es posible generar series artificiales con un exponente de correlación deseado (controlado por α_{sign}) independientemente de las correlaciones presentes en el módulo. Debido a que las correlaciones en el módulo están asociadas a propiedades no lineales de la señal (Kalisky et al., 2005) esto implica que podemos controlar por separado las propiedades lineales y no lineales de la serie compuesta (ver Sec. 2.3).

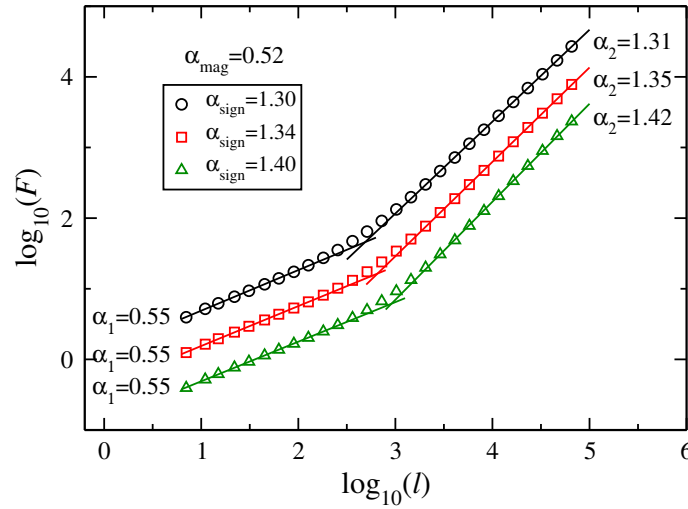


Figura 2.7: Ejemplos de series compuestas generadas mediante la multiplicación de módulos y signos procedentes de distintas series para el caso $\alpha_{\text{mag}} < 1$ y $\alpha_{\text{sign}} > 1$. Obtenemos un crossover que divide el rango en dos regímenes con diferente scaling: $\alpha_1 \simeq \alpha_{\text{mag}}$ para $\ell < \ell_c$ y $\alpha_2 \simeq \alpha_{\text{sign}}$ para $\ell > \ell_c$. El tamaño de serie es 2^{20} y los resultados están promediados sobre 200 experimentos. Las diferentes curvas $F(\ell)$ han sido desplazadas verticalmente para una mejor visualización.

2.2.2. Caso $\alpha_{\text{mag}} < 1, \alpha_{\text{sign}} > 1$

Para este caso observamos comportamientos diferentes a corta y larga escala (Fig. 2.7). Mientras que a corta escala el exponente de correlación $\alpha_1 \simeq \alpha_{\text{mag}}$, a larga escala el signo controla las correlaciones y obtenemos $\alpha_2 \simeq \alpha_{\text{sign}}$. El porqué de este crossover⁵ se puede explicar de la siguiente forma: Teniendo en cuenta que $\alpha_{\text{sign}} > 1$, en ventanas de tamaño pequeño la serie es muy persistente y por tanto, es poco probable que haya un cambio en el signo por lo que estará formada por grandes regiones solo de 1's o -1's (Carretero-Campos et al., 2012), por lo que las fluctuaciones a corta escala dependen únicamente de la serie módulo. Por otro lado, para escalas suficientemente grandes, los cambios de signo dentro de la ventana crearán fluctuaciones mucho más grandes que las creadas debido a las correlaciones del módulo, por tanto, el exponente de correlación tomará valores próximos a α_{sign} .

La posición del crossover ℓ_c entre los dos regímenes depende del tamaño de la secuencia N y puede ser determinada analíticamente considerando que la transición entre estas dos regiones debe ocurrir a una escala ℓ_c para la que las fluctuaciones debidas tanto al módulo como al signo tengan la misma contribución.

Una serie con correlaciones de largo alcance y $\alpha < 1$ es estacionaria, por lo que podemos escribir que las fluctuaciones del módulo a escala ℓ :

⁵El término *crossover* denota un cambio de pendiente en $F(\ell)$

$$F_{\text{mag}}(\ell) = A_{\text{mag}} \ell^{\alpha_{\text{mag}}} \quad (2.19)$$

donde A_{mag} es una constante. Sin embargo, para $\alpha \geq 1$ la serie es no estacionaria y las fluctuaciones a una determinada escala también dependen del tamaño de secuencia N . En particular la varianza crece como $N^{\alpha-1}$, y debido a que generamos series con media cero y varianza unidad, tenemos un factor de normalización $1/N^{\alpha-1}$. De acuerdo con esto, en el Apénd. A demostramos que las fluctuaciones en la serie signo cuando $\alpha \geq 1$ pueden ser escritas como:

$$F_{\text{sign}}(\ell) = B_{\text{sign}} \frac{\ell^{\alpha_{\text{sign}}}}{N^{\alpha_{\text{sign}}-1}} \quad (2.20)$$

donde B_{sign} es una constante. Claramente, la potencia positiva de N ($\alpha_{\text{sign}} - 1 > 0$) en el divisor de (2.20) hace que $F_{\text{sign}}(\ell) < F_{\text{mag}}(\ell)$ a corta escala mientras que $F_{\text{sign}}(\ell) > F_{\text{mag}}(\ell)$ para largas escalas. Este comportamiento también justifica el hecho, comentado previamente, de que a corta escala la serie escala con $\alpha_1 \simeq \alpha_{\text{mag}}$ y para largas escalas lo hace con $\alpha_2 \simeq \alpha_{\text{sign}}$. Por lo tanto, el crossover estará localizado en un punto ℓ_c donde se igualen las fluctuaciones del módulo y signo (Eqs. 2.19 y 2.20) y entonces:

$$\ell_c = \frac{A_{\text{mag}}}{B_{\text{sign}}} N^{\frac{\alpha_{\text{sign}}-1}{\alpha_{\text{sign}}-\alpha_{\text{mag}}}} \propto N^k \quad (2.21)$$

donde

$$k = \frac{\alpha_{\text{sign}} - 1}{\alpha_{\text{sign}} - \alpha_{\text{mag}}} \quad (2.22)$$

Los resultados analíticos obtenidos en Eqs. 2.21 y 2.22 muestran una buena concordancia con las simulaciones de la Fig. 2.8. Vale la pena mencionar que $k < 1$, teniendo en cuenta que $\alpha_{\text{mag}} < 1$. Esto significa que la posición del crossover ℓ_c crece de manera más lenta que el tamaño de la secuencia, y consecuentemente, el crossover será siempre visible para secuencias de tamaño suficientemente grande.

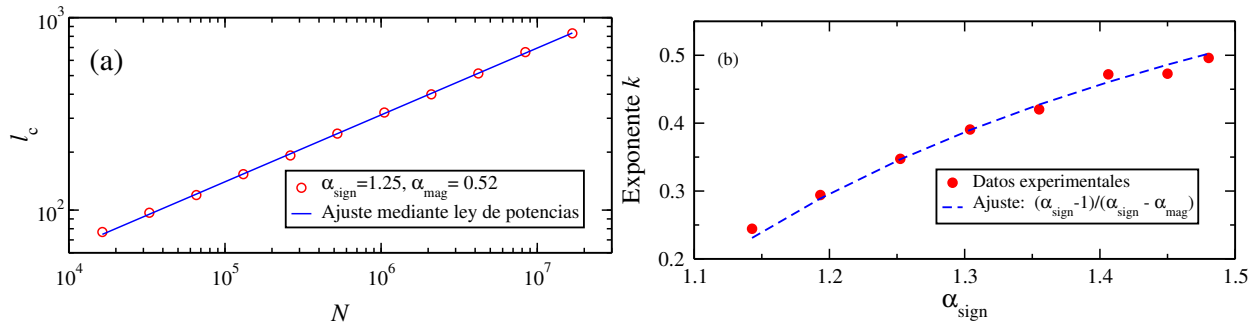


Figura 2.8: Posición del crossover ℓ_c en función del tamaño de la secuencia N . (a) Para comprobar que ℓ_c crece como una ley de potencias en función del tamaño de la secuencia N (Eq. 2.21), generamos series con $\alpha_{\text{sign}} = 1.25$ y $\alpha_{\text{mag}} = 0.52$ y tamaños de secuencia en el rango $[2^{13}, 2^{24}]$. Para cada tamaño obtenemos $F(\ell)$, promediada sobre 1000 experimentos y determinamos la posición del crossover ℓ_c ajustando $F(\ell)$ a la derivada de una sigmoide (Carpena et al., 2002). Este ajuste también nos facilita los valores de la pendiente para los dos regímenes. El ajuste de la curva ℓ_c vs. N (círculos rojos y abiertos) a una ley de potencias (línea sólida azul) ofrece un exponente $k = 0.345$ muy próximo al valor 0.342 predicho por la Eq. 2.22. (b) Repetimos este experimento para diferentes valores de α_{sign} en el rango $[1, 1.5]$ y obtenemos el valor del exponente k para cada uno de ellos. Finalmente, ajustamos la curva de k vs. α_{sign} (círculos rojos y cerrados) a la Eq. 2.22 (línea azul discontinua). El valor obtenido para $\alpha_{\text{mag}} = 0.52$ coincide con el valor utilizado para las simulaciones.

2.2.3. Caso $\alpha_{\text{sign}} > 1, \alpha_{\text{mag}} > 1$

En este caso $F(\ell)$ puede también presentar un crossover aunque será difícil de observar. Para mejorar la comprensión del comportamiento de $F(\ell)$ en esta región seguimos un procedimiento similar al descrito en el caso anterior. De manera análoga a la Subsec. 2.2.2, $\alpha_{\text{sign}} > 1$ y obtenemos para las fluctuaciones en el signo:

$$F_{\text{sign}}(\ell) = B_{\text{sign}} \frac{\ell^{\alpha_{\text{sign}}}}{N^{\alpha_{\text{sign}}-1}} \quad (2.23)$$

Además, para este caso, la serie módulo es ahora no estacionaria ($\alpha_{\text{mag}} > 1$) y de la definición del Movimiento Fraccionario Browniano se sabe que la varianza de la serie crece con $N^{2(\alpha_{\text{mag}}-1)}$. Esto significa que, para mantener la serie con desviación estándar unidad, el método de generación de series (FFM) realiza una división en la serie por un factor $N^{\alpha_{\text{mag}}-1}$ (Kantelhardt et al., 2000) y, por lo tanto, para las fluctuaciones de la serie módulo obtendremos:

$$F_{\text{mag}}(\ell) = B_{\text{mag}} \frac{\ell^{\alpha_{\text{mag}}}}{N^{\alpha_{\text{mag}}-1}} \quad (2.24)$$

De nuevo, la posición del crossover vendrá dada por el valor donde las fluctuaciones del módulo y del signo tengan la misma contribución:

$$B_{\text{mag}} \frac{\ell_c^{\alpha_{\text{mag}}}}{N^{\alpha_{\text{mag}}-1}} = B_{\text{sign}} \frac{\ell_c^{\alpha_{\text{sign}}}}{N^{\alpha_{\text{sign}}-1}} \quad (2.25)$$

$$\ell_c = N \left(\frac{B_{\text{mag}}}{B_{\text{sign}}} \right)^{\frac{1}{\alpha_{\text{sign}} - \alpha_{\text{mag}}}} \propto N \quad (2.26)$$

Esto significa que la posición del crossover crece proporcionalmente al tamaño de la serie. Aquí es importante señalar que la normalización descrita previamente produce una reducción de las fluctuaciones a corta escala (Kantelhardt et al., 2000). Esta reducción se hace más evidente a medida que α aumenta y, por lo tanto, a corta escala ($\ell \ll \ell_c$), las fluctuaciones están controladas por el valor más pequeño $\alpha_1 = \min\{\alpha_{\text{mag}}, \alpha_{\text{sign}}\}$, mientras que a larga escala ($\ell \gg \ell_c$) el exponente de correlación vendrá dado por $\alpha_2 = \max\{\alpha_{\text{mag}}, \alpha_{\text{sign}}\}$. Para calcular los valores de la escala donde tiene lugar el crossover ℓ_c , hemos generado pares de series módulo y signo correlacionadas con $\alpha_{\text{mag}} \in (1, 2)$ y $\alpha_{\text{sign}} \in (1, 1.5)$ y, para cada par de series módulo y signo obtenemos $\alpha_{\text{mag}}, \alpha_{\text{sign}}, B_{\text{mag}}$ y B_{sign} y evaluamos ℓ_c usando la Eq. 2.26. Encontramos dos regiones diferentes:

- i) $\alpha_{\text{mag}} \geq \alpha_{\text{sign}}$. Para este caso, en todos los experimentos obtenemos $\ell_c > N$ por lo que no es posible alcanzar el crossover.
- ii) $\alpha_{\text{mag}} < \alpha_{\text{sign}}$. Aquí, en algunas situaciones obtenemos $\ell_c \leq N$ aunque los valores de $F_{\text{mag}}(\ell)$ y $F_{\text{sign}}(\ell)$ son muy próximos y, por tanto, dificultan la apreciación de un claro crossover. Además, únicamente valores de $\ell_c \leq N/10$ pueden ser observados ya que el DFA coge un tamaño máximo de ventana $N/10$ (Hu, 2001).

En definitiva, raramente observamos crossovers dentro de esta región y la serie compuesta mostrará un único scaling en todo el rango de escalas, siendo el exponente de correlación $\alpha = \alpha_1 \simeq \min\{\alpha_{\text{mag}}, \alpha_{\text{sign}}\}$.

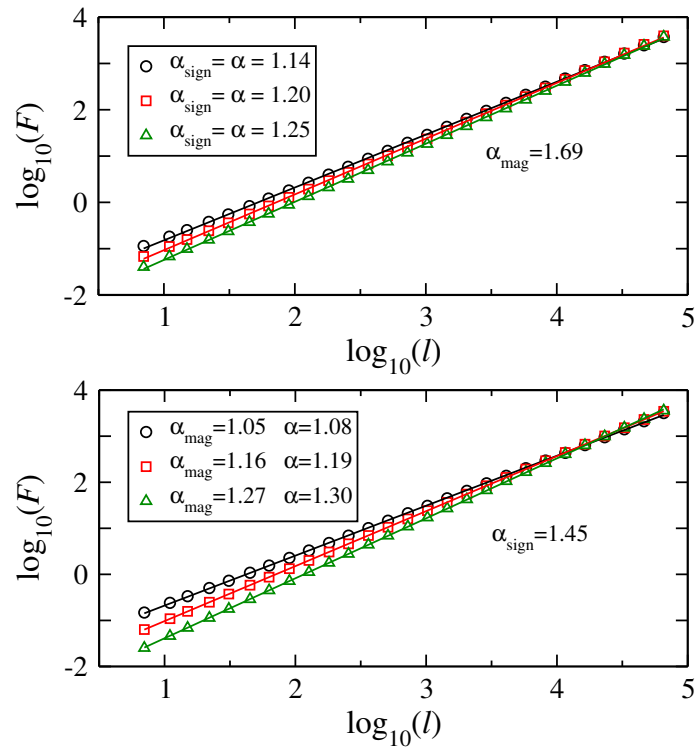


Figura 2.9: Ejemplo de series compuestas generadas a través de la multiplicación de módulos y signos procedentes de series independientes para $\alpha_{\text{mag}} > 1$ y $\alpha_{\text{sign}} > 1$. En esta región no observamos crossover (ver texto). (a) Ejemplo con $\alpha_{\text{sign}} < \alpha_{\text{mag}}$. El exponente de correlación de la serie compuesta toma el valor α_{sign} ($\min\{\alpha_{\text{mag}}, \alpha_{\text{sign}}\}$) para cada serie. (b) Ejemplo con $\alpha_{\text{sign}} > \alpha_{\text{mag}}$. Las correlaciones en la serie compuesta están principalmente controladas por el módulo, aunque el exponente de correlación es ligeramente más alto que el del módulo. Este efecto se hace más notable a medida que la diferencia entre el exponente de correlación del módulo y del signo disminuye, ya que el crossover es menos pronunciado y, por tanto, existe una pequeña contribución debida al segundo régimen (situado después del crossover).

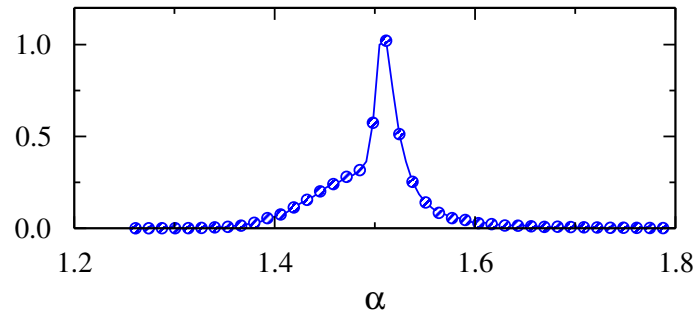


Figura 2.10: Distribución de exponentes para series obtenidas con $\alpha_{\text{mag}} \simeq 1.9$ y $\alpha_{\text{sign}} \simeq 1.5$. El tamaño de la serie es 2^{20} y los resultados son fruto del promedio sobre 10^5 series.

Otra conclusión que podemos extraer de este caso es el hecho de que no es posible generar series artificiales con un exponente de correlación $\alpha > 1.5$ mediante la composición de módulos y signos procedentes de series originales independientes. Para corroborar esto, intentamos obtener el máximo exponente de correlación posible a través de composición de series, usando para ello $\alpha_{\text{in}} = 2$ para el módulo ($\alpha_{\text{mag}} \simeq 1.9$) y $\alpha_{\text{in}} = 2$ para el signo ($\alpha_{\text{sign}} \simeq 1.5$) (ver Fig. 2.3). Multiplicamos cada par de series módulo y signo generadas, calculamos α para la serie compuesta y representamos la distribución de valores. Los resultados (ver Fig. 2.10) muestran una distribución con un pico muy pronunciado en $\alpha = 1.5$ ($\min\{\alpha_{\text{mag}}, \alpha_{\text{sign}}\}$), en concordancia con lo que hemos explicado previamente. Si bien es cierta la existencia de algunas series con $\alpha > 1.5$, estas series están asociadas a situaciones donde el scaling de la serie compuesta no es bueno, junto con las pocas veces en las que el crossover puede ser observable.

Finalmente, la tabla 2.1 resume los resultados obtenidos en esta sección.

α_{sign}	α_{mag}	α	CROSSOVER
< 1	$[0.5, 2]$	α_{sign}	no
> 1	< 1	$\ell < \ell_c \quad \alpha_1 = \alpha_{\text{mag}}$	$\ell_c \propto N^k$
		$\ell > \ell_c \quad \alpha_2 = \alpha_{\text{sign}}$	$k = \frac{\alpha_{\text{sign}} - 1}{\alpha_{\text{sign}} - \alpha_{\text{mag}}}$
> 1	> 1	$\min\{\alpha_{\text{mag}}, \alpha_{\text{sign}}\}$	no observable

Tabla 2.1: Resultados obtenidos para la composición de series a través de módulos y signos independientes.

2.3. Propiedades multifractales de series compuestas

En la sección anterior nos hemos limitado únicamente a estudiar las correlaciones lineales en las series compuestas. Sin embargo, sabemos que series con módulos correlacionados ($\alpha_{\text{mag}} > 0.5$) y signos no correlacionados ($\alpha_{\text{sign}} = 0.5$) también poseen correlaciones no lineales (propiedades multifractales) (Ashkenazy, 2003; Kalisky et al., 2005). Por tanto, en esta sección vamos a analizar las propiedades multifractales de series compuestas a través de módulos y signos independientes.

Los resultados obtenidos para series compuestas en la Sec. 2.2 indican la existencia de un crossover a escala ℓ_c siempre y cuando $\alpha_{\text{sign}} > 1$. Dicho comportamiento podría indicar la existencia de dos espectros multifractales diferentes, uno por encima y otro por debajo de ℓ_c . Además, no podemos garantizar que $\ell_c = \text{cte}$ para los diferentes momentos de orden q , dificultando por tanto un cálculo directo de ambos espectros multifractales. Por esta razón, nos restringimos a la región $\alpha_{\text{sign}} < 1$ donde, en primer lugar la serie compuesta tiene un único scaling y, en segundo lugar las correlaciones lineales en dicha serie están directamente controladas por el signo ($\alpha = \alpha_{\text{sign}}$) (ver Sec. 2.1). También se debe tener en cuenta que a la hora de calcular el espectro multifractal de series compuestas existen inestabilidades numéricas cuando el exponente de correlación del módulo toma valores superiores a $\alpha_{\text{mag}} > 1.2$. Por todo esto, vamos a estudiar las propiedades multifractales de series compuestas con α_{mag} y α_{sign} en los intervalos $[0.5, 1.2]$ y $[0.5, 1]$ respectivamente. Específicamente, calculamos el espectro multifractal para cada serie compuesta usando MFDFA (ver Sec. 1.5) y estudiamos sistemáticamente dos propiedades: la anchura del espectro multifractal, $\Delta\zeta$, y la localización de su centro, ζ_{max} . En cuanto a las propiedades del espectro multifractal, observamos que la anchura $\Delta\zeta$ depende únicamente de α_{mag} , y es prácticamente independiente del valor de α_{sign} . Ambas propiedades se muestran en la Fig. 2.11 a), donde también nos percatamos de que la dependencia de $\Delta\zeta$ con α_{mag} es prácticamente lineal. Esta propiedad es muy interesante, ya que sabiendo la correlación del módulo usado para la composición α_{mag} , podemos controlar directamente las propiedades no lineales de la serie, ya que vienen cuantificadas por la anchura del espectro $\Delta\zeta$.

También estudiamos cómo las correlaciones lineales presentes en la serie compuesta, las cuales son controladas por α_{sign} ($\alpha = \alpha_{\text{sign}}$), están implicadas en la localización del centro del espectro multifractal ζ_{max} . Observamos que, dado un valor fijo α_{mag} , (y por tanto $\Delta\zeta$) el espectro multifractal es desplazado proporcionalmente al valor que toma el exponente de correlación del signo α_{sign} (ver Fig. 2.11 b). De hecho, si calculamos numéricamente la localización del centro del espectro, ζ_{max} , obtenemos una dependencia lineal muy buena de ζ_{max} con α_{sign} (ver Fig. 2.11 c), con pendiente $\simeq 1$.

A modo de conclusión, las propiedades multifractales de series temporales obtenidas

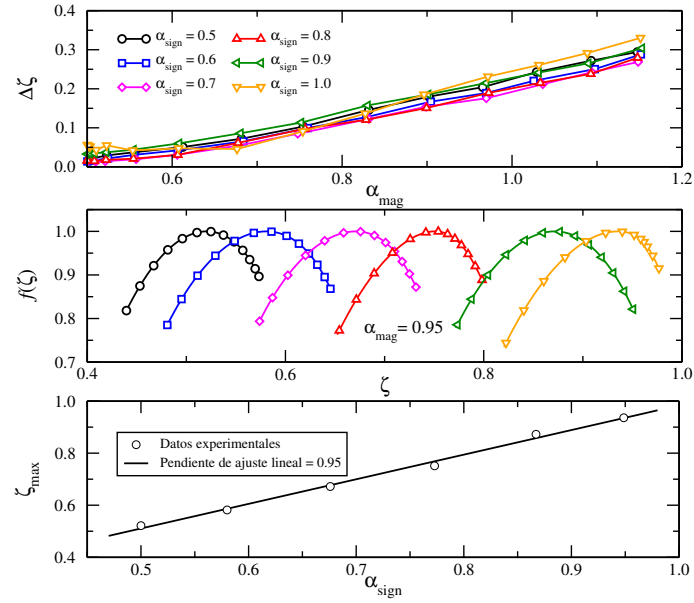


Figura 2.11: (a) Relación entre $\Delta\zeta$ y α_{mag} de series compuestas mediante la multiplicación de módulos y signos independientes. $\Delta\zeta$ aumenta de forma lineal con α_{mag} cuando la serie módulo abandona la región en la que no presenta correlaciones ($\alpha_{\text{mag}} > 0.5$), mientras que α_{sign} no juega un papel relevante en el valor de $\Delta\zeta$. Las series tienen un tamaño de 2^{18} y los resultados están promediados sobre un conjunto de 50 series. (b) Ejemplo de espectros multifractales de una serie de tamaño 2^{18} con $\alpha_{\text{mag}} = 0.95$ y α_{sign} en el intervalo $[0.5, 1]$. Para obtener el espectro multifractal, el MF DFA (1.21) se ha ejecutado utilizando momentos comprendidos en el intervalo $q \in [-5, 5]$. A pesar de variar α_{sign} , $\Delta\zeta$ toma prácticamente el mismo valor para todos los casos. Los espectros multifractales están centrados en $\zeta_{\text{max}} \simeq \alpha_{\text{sign}}$. (c) Relación entre α_{sign} y ζ_{max} para los espectros multifractales obtenidos en (b). Los datos son ajustados mediante regresión lineal con pendiente 0.95.

mediante la multiplicación de módulos y signos independientes están totalmente controladas por dos parámetros: los exponentes de correlación α_{sign} y α_{mag} . Mientras que el primero controla las correlaciones lineales de la serie compuesta (α) y la localización del centro del espectro multifractal (ζ_{max}), el segundo cuantifica la anchura del espectro multifractal ($\Delta\zeta$) y, por tanto, la intensidad de las no linealidades. Obviamente, este procedimiento puede ser usado como un algoritmo de generación de series artificiales en las que, además de controlar las correlaciones lineales (FFM), también podamos controlar las propiedades multifractales.

Capítulo 3

Función de autocorrelación de series módulo y signo de fGns

En el Cap. 2 hemos utilizado sistemáticamente el DFA para la obtención de exponentes de correlación de señales módulo y signo. Sin embargo, el DFA es un método de medida de correlación indirecto (Peng et al., 1994b), que como veremos únicamente ofrece resultados interpretables cuando las correlaciones de la serie bajo estudio decaen mediante una ley de potencias. Uno de los mayores problemas que plantea el DFA a la hora de utilizarlo en datos procedentes de sistemas reales es que en la mayoría de los casos la curva $F(\ell)$ vs. ℓ no se puede ajustar mediante una ley de potencias i.e. no hay scaling, por lo que no tendría sentido alguno hablar del exponente de correlación α medido por el DFA, ya que no está claro qué significan las desviaciones sobre la ley de potencias. Es aquí donde reside nuestro interés en el estudio de la propia función de autocorrelación, que a pesar de ser en muchos casos ruidosa y sensible al tamaño N de la serie temporal, nos facilita la información exacta que estamos buscando sin importar si la propia función de autocorrelación, y por consiguiente la curva $F(\ell)$ vs. ℓ se ajusta de manera óptima o no mediante una ley de potencias. Además, desde el descubrimiento del ruido $1/f$ se ha puesto el foco únicamente en las correlaciones que decaen mediante una ley de potencias, obviando en la mayoría de los casos que la función de autocorrelación de la señal revela propiedades importantes sobre el sistema independientemente de su dependencia funcional, por lo que también consideramos interesante el estudio de la función de autocorrelación cuando esta no decae mediante una ley de potencias.

De nuevo, de forma análoga a lo que hicimos en el capítulo anterior, estamos interesados en el estudio de las correlaciones de la serie módulo y la serie signo, dada la importancia que tienen a la hora de entender la dinámica de sistemas complejos. Para abordarlo, en primer lugar calculamos de manera analítica la expresión que relaciona la

función de autocorrelación de la serie módulo $C_{|x|}$ con la función de autocorrelación de la serie original C_x para ruidos lineales Gaussianos, i.e. series temporales en las que los valores siguen una distribución Gaussiana, y de acuerdo con la definición de Schreiber y Schmitz (Schreiber and Schmitz, 2000), las fases son aleatorias (ver Cap. 4). También calculamos la relación entre la función de autocorrelación de la serie cuadrática C_{x^2} en función de la función de autocorrelación de la serie original C_x y probamos que la relación analítica entre la función de autocorrelación del signo C_s y la de la serie original C_x (Apostolov et al., 2008) no sólo es válida para distribuciones normales, sino que lo es para todas las distribuciones de valores simétricas. Después, discutimos la relación entre la función de autocorrelación del signo C_s y la del módulo $C_{|x|}$ en fGns y por último establecemos un método para la cuantificación de la no-linealidad en series temporales a partir del estudio de la función de autocorrelación de series no lineales generadas a través de la composición de módulos y signos independientes (ver Sec. 2.2).

3.1. Autocorrelación de la serie módulo para ruidos fraccionarios Gaussianos

Dado un ruido fraccionario Gaussiano $\{x_i\}$ cuya distribución de valores corresponde a una Gaussiana de media nula y desviación estándar unidad $\mathcal{N}(0, 1)$, pretendemos calcular la función de autocorrelación de su serie módulo $\{|x_i|\}$, $C_{|x|}(\ell)$ en función de la autocorrelación de la serie de partida (a lo largo del texto nos referimos a esta serie de manera reiterada como *serie original*) $\{x_i\}$, $C_x(\ell)$.

Consideramos que la serie $\{x_i\}$, $i \in \{1, \dots, N\}$ está compuesta por variables Gaussianas con $\bar{x} = 0$ y $\sigma_x = 1$. Si solo consideramos las correlaciones lineales presentes en la serie, la densidad de probabilidad conjunta para dos variables separadas temporalmente por ℓ ($x_i, x_{i+\ell}$) vendrá dada por la distribución Gaussiana bivariada (Johnson et al., 2002):

$$\rho_{x,\ell}(\xi, \varphi) \equiv \text{Prob} \{x_i = \xi, x_{i+\ell} = \varphi\} = \frac{1}{2\pi\sqrt{1 - k_x^2(\ell)}} \exp \left\{ -\frac{\xi^2 + \varphi^2 - 2k_x(\ell)\xi\varphi}{2[1 - k_x^2(\ell)]} \right\}, \quad (3.1)$$

donde $k_x(\ell)$ es la autocovarianza (también coincide con la autocorrelación por el hecho de ser variables con media nula y desviación estándar unidad) de la serie $\{x_i\}$ a distancia ℓ :

$$k_x(\ell) = \langle \xi\varphi \rangle. \quad (3.2)$$

Por otro lado, denotando $A_k = \sqrt{1 - k_x^2(\ell)}$, la distribución de probabilidad conjunta de

la serie módulo puede ser obtenida como:

$$\begin{aligned} F_{|x|,\ell}(\xi, \varphi) &\equiv \text{Prob} \{ |x_i| \leq \xi, |x_{i+\ell}| \leq \varphi \} = \int_{-\xi}^{\xi} d\xi' \int_{-\varphi}^{\varphi} d\varphi' \rho_{x,\ell}(\xi', \varphi') \\ &= \int_{-\xi}^{\xi} d\xi' \int_{-\varphi}^{\varphi} d\varphi' \frac{2}{\pi A_k} \exp \left(-\frac{\xi'^2 + \varphi'^2 - 2k_x(\ell)\xi'\varphi'}{2A_k^2} \right) = I. \end{aligned} \quad (3.3)$$

A continuación procedemos a la resolución de la integral I :

$$\begin{aligned} I &= \frac{1}{2\pi A_k} \left[\int_{-\xi}^{\xi} d\xi' \left(\int_{-\varphi}^{\varphi} d\varphi' \exp \left(-\frac{\xi'^2 + \varphi'^2 - 2k_x(\ell)\xi'\varphi'}{2A_k^2} \right) \right) \right. \\ &\quad \left. + \int_0^{\varphi} d\varphi' \exp \left(-\frac{\xi'^2 + \varphi'^2 - 2k_x(\ell)\xi'\varphi'}{2A_k^2} \right) \right] \end{aligned} \quad (3.4)$$

$$\begin{aligned} &= \frac{1}{2\pi A_k} \left[\int_{-\xi}^0 d\xi' \int_{-\varphi}^0 d\varphi' \exp \left(-\frac{\xi'^2 + \varphi'^2 - 2k_x(\ell)\xi'\varphi'}{2A_k^2} \right) \right. \\ &\quad + \int_0^{\xi} d\xi' \int_{-\varphi}^0 d\varphi' \exp \left(-\frac{\xi'^2 + \varphi'^2 - 2k_x(\ell)\xi'\varphi'}{2A_k^2} \right) \\ &\quad + \int_{-\xi}^0 d\xi' \int_0^{\varphi} d\varphi' \exp \left(-\frac{\xi'^2 + \varphi'^2 - 2k_x(\ell)\xi'\varphi'}{2A_k^2} \right) \\ &\quad \left. + \int_0^{\xi} d\xi' \int_0^{\varphi} d\varphi' \exp \left(-\frac{\xi'^2 + \varphi'^2 - 2k_x(\ell)\xi'\varphi'}{2A_k^2} \right) \right], \end{aligned} \quad (3.5)$$

En la Eq. 3.5, desarrollamos las exponenciales y manipulamos para obtener los mismos límites de integración en todos los sumandos:

$$\begin{aligned} I &= \frac{1}{2\pi A_k} \left[\int_0^{\xi} d\xi' \int_0^{\varphi} d\varphi' \exp \left(-\frac{\xi'^2 + \varphi'^2}{2A_k^2} \right) \exp \left(\frac{k_x(\ell)\xi'\varphi'}{2A_k^2} \right) \right. \\ &\quad + \int_0^{\xi} d\xi' \int_0^{\varphi} d\varphi' \exp \left(-\frac{\xi'^2 + \varphi'^2}{2A_k^2} \right) \exp \left(-\frac{k_x(\ell)\xi'\varphi'}{2A_k^2} \right) \\ &\quad + \int_0^{\xi} d\xi' \int_0^{\varphi} d\varphi' \exp \left(-\frac{\xi'^2 + \varphi'^2}{2A_k^2} \right) \exp \left(-\frac{k_x(\ell)\xi'\varphi'}{2A_k^2} \right) \\ &\quad \left. + \int_0^{\xi} d\xi' \int_0^{\varphi} d\varphi' \exp \left(-\frac{\xi'^2 + \varphi'^2}{2A_k^2} \right) \exp \left(\frac{k_x(\ell)\xi'\varphi'}{2A_k^2} \right) \right] \\ &= \frac{1}{2\pi A_k} \int_0^{\xi} d\xi' \int_0^{\varphi} d\varphi' \exp \left(-\frac{\xi'^2 + \varphi'^2}{2A_k^2} \right) \left[2 \exp \left(\frac{k_x(\ell)\xi'\varphi'}{2A_k^2} \right) + 2 \exp \left(-\frac{k_x(\ell)\xi'\varphi'}{2A_k^2} \right) \right], \end{aligned} \quad (3.6)$$

por lo que finalmente podemos expresar el valor de $F_{|x|,\ell}(\xi, \varphi)$ como:

$$\begin{aligned} F_{|x|,\ell}(\xi, \varphi) &\equiv \text{Prob} \{ |x_i| \leq \xi, |x_{i+\ell}| \leq \varphi \} = I \\ &= \frac{2}{\pi A_k} \int_0^\xi d\xi' \int_0^\varphi d\varphi' \exp \left(-\frac{\xi'^2 + \varphi'^2}{2A_k^2} \right) \cosh \left(\frac{k_x(\ell)\xi'\varphi'}{2A_k^2} \right). \end{aligned} \quad (3.7)$$

La densidad de probabilidad vendrá dada por:

$$\rho_{|x|,\ell}(\xi, \varphi) = \frac{\partial^2}{\partial \xi \partial \varphi} F_{|x|,\ell}(\xi, \varphi) = \frac{2}{\pi A_k} \exp \left(-\frac{\xi'^2 + \varphi'^2}{2A_k^2} \right) \cosh \left(\frac{k_x(\ell)\xi'\varphi'}{2A_k^2} \right), \quad (3.8)$$

y la autocovarianza del módulo:

$$k_{|x|}(\ell) = \langle |x_i| \cdot |x_{i+\ell}| \rangle = \int_0^\infty d\xi' \int_0^\infty d\varphi' \rho_{|x|,\ell}(\xi', \varphi'). \quad (3.9)$$

Tomando el cambio de variable $\xi = \frac{\xi'}{A_k}$ y $\varphi = \frac{\varphi'}{A_k}$:

$$\begin{aligned} k_{|x|}(\ell) &= \frac{2}{\pi A_k} \int_0^\infty A_k^2 \xi d\xi' \int_0^\infty A_k^2 \varphi d\varphi' \exp \left(-\frac{(\xi A_k)^2 + (\varphi A_k)^2}{2A_k^2} \right) \cosh(k_x(\ell)\xi\varphi) \\ &= \frac{2A_k^3}{\pi} \int_0^\infty \xi \exp \left(-\frac{\xi^2}{2} \right) d\xi \int_0^\infty \varphi \exp \left(-\frac{\varphi^2}{2} \right) \cosh(k_x(\ell)\xi\varphi) d\varphi. \end{aligned} \quad (3.10)$$

Llamando I_1 a:

$$I_1 = \int_0^\infty \varphi \exp \left(-\frac{\varphi^2}{2} \right) \cosh(k_x(\ell)\xi\varphi) d\varphi, \quad (3.11)$$

y aplicando la técnica de separación de variables podemos expresar I_1 como:

$$I_1 = \frac{1}{\xi} \frac{\partial}{\partial k_x(\ell)} \left[\int_0^\infty \exp \left(-\frac{\varphi^2}{2} \right) \sinh(k_x(\ell)\xi\varphi) d\varphi \right]. \quad (3.12)$$

Ahora, resolvemos la integral que aparece en (3.12), denotándola por I_2 :

$$\begin{aligned} I_2 &= \int_0^\infty \exp \left(-\frac{\varphi^2}{2} \right) \sinh(k_x(\ell)\xi\varphi) d\varphi \\ &= \int_0^\infty d\varphi \exp \left(-\frac{\varphi^2}{2} \right) \frac{e^{k_x(\ell)\xi\varphi} - e^{-k_x(\ell)\xi\varphi}}{2} \\ &= \frac{1}{2} \left[\underbrace{\int_0^\infty d\varphi e^{-\varphi^2/2 + k_x(\ell)\xi\varphi}}_{I_{2+}} - \underbrace{\int_0^\infty d\varphi e^{-\varphi^2/2 - k_x(\ell)\xi\varphi}}_{I_{2-}} \right]. \end{aligned} \quad (3.13)$$

Utilizando la siguiente identidad (Gradshteyn and Ryzhik, 2014):

$$\int_0^\infty e^{-a^2x^2+bx} dx = \frac{\sqrt{\pi}}{2a} \exp\left(\frac{b^2}{4a^2}\right) \left[1 + \operatorname{erf}\left(\frac{b}{2a}\right)\right], \quad (3.14)$$

podemos expresar I_{2+} y I_{2-} como:

$$I_{2+} \equiv \{a = \frac{1}{\sqrt{2}}, b = k_x(\ell)\xi\} = \frac{\sqrt{2\pi}}{2} \exp\left(\frac{k_x^2(\ell)\xi^2}{2}\right) \left[1 + \operatorname{erf}\left(\frac{k_x(\ell)\xi}{\sqrt{2}}\right)\right] \quad (3.15)$$

$$I_{2-} \equiv \{a = \frac{1}{\sqrt{2}}, b = k_x(\ell)\xi\} = \frac{\sqrt{\pi}}{2} \exp\left(\frac{k_x^2(\ell)\xi^2}{2}\right) \left[1 + \operatorname{erf}\left(-\frac{k_x(\ell)\xi}{\sqrt{2}}\right)\right]. \quad (3.16)$$

Operamos y haciendo uso del carácter impar de la función error:

$$\operatorname{erf}(x) = \frac{2}{\pi} \int_0^x e^{-t^2} dt, \quad (3.17)$$

ya estamos en disposición de resolver la integral I_2 (3.13)

$$I_2 = \int_0^\infty \exp\left(-\frac{\varphi^2}{2}\right) \sinh(k_x(\ell)\xi\varphi) d\varphi = \sqrt{\frac{\pi}{2}} \exp\left(\frac{k_x^2(\ell)\xi^2}{2}\right) \operatorname{erf}\left[\frac{k_x(\ell)\xi}{\sqrt{2}}\right], \quad (3.18)$$

lo que nos permite solucionar (3.11)

$$\begin{aligned} \int_0^\infty \varphi \exp\left(-\frac{\varphi^2}{2}\right) \cosh(k_x(\ell)\xi\varphi) d\varphi &= \frac{1}{\xi} \frac{\partial}{\partial k_x(\ell)} \left[\sqrt{\frac{\pi}{2}} \exp\left(\frac{k_x^2(\ell)\xi^2}{2}\right) \operatorname{erf}\left(\frac{k_x(\ell)\xi}{\sqrt{2}}\right) \right] \\ &= \sqrt{\frac{\pi}{2}} k_x(\ell)\xi \exp\left(\frac{k_x^2(\ell)\xi^2}{2}\right) \operatorname{erf}\left(\frac{k_x(\ell)\xi}{\sqrt{2}}\right) + 1. \end{aligned} \quad (3.19)$$

Ahora, al sustituir (3.19) en (3.10) podemos escribir la autocovarianza de la serie módulo como:

$$\begin{aligned} k_{|x|}(\ell) &= \frac{2A_k^3}{\pi} \int_0^\infty \xi \exp\left(-\frac{\xi^2}{2}\right) d\xi \left[\sqrt{\frac{\pi}{2}} k_x(\ell)\xi \exp\left(\frac{k_x^2(\ell)\xi^2}{2}\right) \operatorname{erf}\left(\frac{k_x(\ell)\xi}{\sqrt{2}}\right) + 1 \right] \\ &= \sqrt{\frac{2}{\pi}} k_x(\ell) A_k^3 \int_0^\infty \xi^2 \exp\left(-\frac{A_k^2 \xi^2}{2}\right) \operatorname{erf}\left(\frac{k_x(\ell)\xi}{\sqrt{2}}\right) d\xi + \frac{2A_k^3}{\pi} \int_0^\infty \xi \exp\left(-\frac{\xi^2}{2}\right) d\xi. \end{aligned} \quad (3.20)$$

Resolvemos la primera de las dos integrales que aparecen en (3.20) (la llamamos I_3):

$$I_3 = \int_0^\infty \xi^2 \exp\left(-\frac{A_k^2 \xi^2}{2}\right) \operatorname{erf}\left(\frac{k_x(\ell)\xi}{\sqrt{2}}\right) d\xi, \quad (3.21)$$

y usamos la identidad extraída de (Ng and Geller, 1969):

$$\int_0^\infty \xi^2 d\xi \exp(-b^2 \xi^2) \operatorname{erf}(a\xi) = \frac{\sqrt{\pi}}{4b^3} \operatorname{sign}(a) - \frac{1}{2\sqrt{\pi}} \left[\frac{1}{b^3} \arctan\left(\frac{b}{a}\right) - \frac{a}{b^2(a^2 + b^2)} \right]$$

$$\text{con } a = \frac{K}{\sqrt{2}}, b = \frac{A_k}{\sqrt{2}}, \quad (3.22)$$

para obtener:

$$\sqrt{\frac{2}{\pi}} k_x(\ell) A_k^3 I_3 = k_x(\ell) - \frac{2}{\pi} k_x(\ell) \tan^{-1} \left(\frac{A_k}{k_x(\ell)} \right) + \frac{2}{\pi} k_x^2(\ell) A_k. \quad (3.23)$$

Para el caso de la segunda integral que aparece en (3.20) (la llamamos I_4):

$$I_4 = \int_0^\infty \xi \exp\left(-\frac{\xi^2}{2}\right) d\xi \quad (3.24)$$

usamos la identidad encontrada en (Ng and Geller, 1969):

$$\int_0^\infty x e^{-ax^2+bx} dx = \frac{1}{2a^2} \left[\frac{\sqrt{\pi}b}{2a} \exp\left(\frac{b^2}{4a^2}\right) \operatorname{erfc}\left(-\frac{b}{2a}\right) + 1 \right]$$

$$\text{con } \{a = \frac{1}{\sqrt{2}}, b = 0\}$$

siendo $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x)$. Obtenemos $I_4 = 1$, por lo que:

$$\sqrt{\frac{2}{\pi}} k_x(\ell) A_k^3 I_4 = \sqrt{\frac{2}{\pi}} k_x(\ell) A_k^3. \quad (3.26)$$

Finalmente, reemplazando (3.23) y (3.26) en (3.20) obtenemos que la autocovarianza del módulo puede ser expresada como:

$$k_{|x|}(\ell) = k_x(\ell) + \frac{2}{\pi} \left[\sqrt{1 - k_x^2(\ell)} - k_x(\ell) \tan^{-1} \left(\frac{A_k}{k_x(\ell)} \right) \right]. \quad (3.27)$$

Ahora bien, utilizando la conocida relación trigonométrica:

$$\tan^2(\alpha) = \frac{1}{\cos^2(\alpha) - 1}, \quad (3.28)$$

siendo para nuestro caso:

$$x = \frac{A_k}{k_x(\ell)} \Rightarrow \tan^{-1} \frac{A_k}{k_x(\ell)} = \cos^{-1} \left(\frac{1}{\sqrt{1 + \frac{A_k^2}{k_x^2(\ell)}}} \right) = \cos^{-1}(k_x(\ell)). \quad (3.29)$$

Finalmente, la relación entre la autocovarianza de la serie módulo y la serie original para un fGn puede expresarse como:

$$k_{|x|}(\ell) = k_x(\ell) + \frac{2}{\pi} \left[\sqrt{1 - k_x^2(\ell)} - k_x(\ell) \cos^{-1}(k_x(\ell)) \right]. \quad (3.30)$$

Por último, para obtener una relación analítica entre la autocorrelación del módulo y la autocorrelación de la señal original debemos tener en cuenta que $\{x_i\}$ únicamente posee correlaciones lineales, por lo que la función de autocorrelación a distancia ℓ vendrá dada por:

$$C_x(\ell) = \frac{\langle x_i x_{i+\ell} \rangle - \langle x_i \rangle \langle x_{i+\ell} \rangle}{\sigma_x^2} = \langle x_i x_{i+\ell} \rangle, \quad (3.31)$$

donde $\langle \cdot \rangle$ denota el promedio sobre la serie y σ_x^2 es la varianza de la serie. Considerando que la serie es estacionaria, de media nula y varianza unidad, el valor de la autocorrelación coincide con el de la autocovarianza y, por tanto:

$$k_x(\ell) \equiv \langle x_i x_{i+\ell} \rangle - \langle x_i \rangle \langle x_{i+\ell} \rangle = C_x(\ell). \quad (3.32)$$

Por otro lado, para la serie módulo tenemos que:

$$\begin{aligned} \langle |x_i| \rangle &= \langle |x_{i+\ell}| \rangle = \sqrt{\frac{2}{\pi}} \\ \sigma_{|x|}^2 &= 1 - \frac{2}{\pi}, \end{aligned} \quad (3.33)$$

por lo que su función de autocorrelación vendrá dada por:

$$C_{|x|}(\ell) = \frac{\langle |x_i| \cdot |x_{i+\ell}| \rangle - \langle |x_i| \rangle \langle |x_{i+\ell}| \rangle}{\sigma_{|x|}^2} = \frac{\pi k_{|x|}(\ell)}{\pi - 2}. \quad (3.34)$$

Reemplazando (3.30) en (3.34) obtenemos la relación analítica entre la autocorrelación

del módulo y la de la señal original:

$$C_{|x|}(\ell) = \frac{2 \left[C_x(\ell) \text{sen}^{-1}(C_x(\ell)) - 1 + \sqrt{1 - C_x^2(\ell)} \right]}{\pi - 2}. \quad (3.35)$$

Es fácil comprobar que (3.35) siempre toma valores positivos, lo que implica que un fGn nunca puede tener anticorrelaciones en su serie módulo (ver Fig. 3.1). Además, si consideramos pequeños valores en la autocorrelación podemos aproximar la Eq. 3.35 mediante un desarrollo en serie de Teylor, obteniendo:

$$C_{|x|}(\ell) = \frac{1}{\pi - 2} C_x^2(\ell) + \mathcal{O}(C_x^4(\ell)) \quad (3.36)$$

Por tanto, para pequeños valores, la autocorrelación del módulo se comporta esencialmente como el cuadrado de la autocorrelación de la serie original. De hecho, en la Fig. 3.1 se puede apreciar que el error en (3.36) está en torno al 2 % para $C_x = \pm 0.5$, lo que nos hace poder considerar en el futuro esta aproximación como correcta para la mayoría de los casos en los que usemos datos reales. También vale la pena mencionar el hecho de que en la Fig. 3.1 veamos dos series con las mismas correlaciones lineales (cuadrados azules y triángulos verdes), ambas poseen correlaciones en el módulo y sin embargo una de ellas está situada sobre la curva de la Eq. 3.35 (cuadrados azules) mientras que la otra (triángulos verdes) está totalmente alejada de dicha curva, dando esto a entender, que el hecho de poseer correlaciones en el módulo no necesariamente indica la existencia de propiedades no-lineales, sino la desviación sobre su expectación en el caso de ruidos lineales (3.35).

Además, teniendo en cuenta la Eq. 1.12 ($C(\ell) \sim \text{sgn}(1 - \gamma)/\ell^\gamma$), válida para ruidos fraccionarios Gaussianos estacionarios ($0 < H \leq 1$) podemos extraer algunas conclusiones. Recordando la relación que une al exponente de autocorrelación γ y el exponente de Hurst H , $H = (2 - \gamma)/2$ concluimos que la serie tendrá correlaciones positivas cuando $\gamma < 1$, ya que en ese caso $H > 0.5$. Los procesos físicos que producen series de salida con estas características son denominados como *procesos con larga memoria* o *largo alcance* (Beran, 1994; Rangarajan and Ding, 2000) ya que la función de autocorrelación decae de forma muy lenta con un exponente $\gamma < 1$. Por el contrario, cuando $\gamma > 1$ y, por tanto, $H < 0.5$ la serie posee correlaciones negativas o anticorrelaciones y aunque la función de autocorrelación también decae mediante una ley de potencias no podríamos hablar de *anticorrelaciones de largo alcance* ya que decaen de una manera muy rápida. Obviamente, esto nos induce a pensar que no es posible obtener ruidos fraccionarios Gaussianos con correlaciones positivas y cuya función de autocorrelación decaiga más rápido que $1/\ell$. Ahora bien, reemplazando (1.12) en (3.36) obtenemos que la serie módulo de ruidos fraccionarios

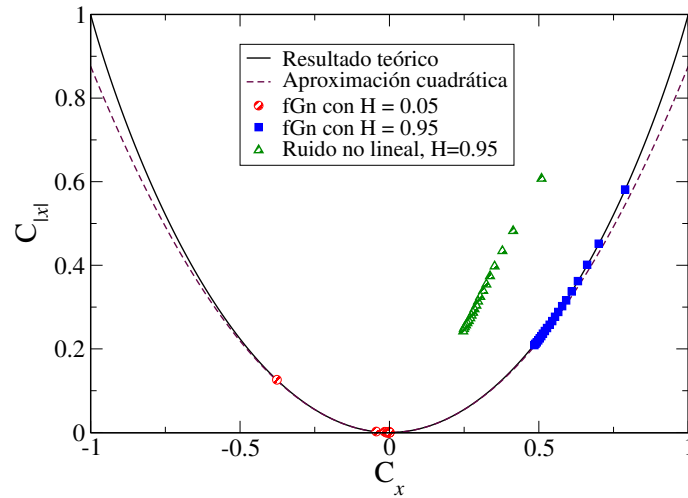


Figura 3.1: Autocorrelación de la serie módulo $C_{|x|}$ en función de la autocorrelación de la serie original C_x . La línea sólida corresponde a la expresión obtenida mediante la Eq. 3.35 mientras que la línea discontinua corresponde a su aproximación cuadrática, Eq. 3.36. Los símbolos representan los valores obtenidos para la autocorrelación del módulo a distancias $\ell = 1, 2, 3, \dots, 20$ para dos ruidos fraccionarios Gaussianos (los círculos corresponden a $H = 0.05$ y los cuadrados a $H = 0.95$) y un ruido no lineal con correlaciones lineales $H = 0.95$ (triángulos verdes). Los fGns han sido generados por el método de filtrado de Fourier mientras que el ruido no lineal ha sido generado mediante la composición de módulos y signos independientes (todas las series tienen un tamaño $N = 2^{20}$). En concreto, el fGn con $H = 0.95$ corresponde a un ruido altamente correlacionado mientras que el fGn con $H = 0.05$ corresponde a un ruido muy anticorrelacionado. El ruido no lineal ($H = 0.95$) muestra un comportamiento totalmente alejado del que predice la Eq. 3.35 para ruidos lineales. En todos los casos, independientemente del valor de la función de autocorrelación de la serie original, las correlaciones en la serie módulo son siempre positivas en los tres ruidos.

Gaussianos estacionarios también decae mediante una ley de potencias:

$$C_{|x|}(\ell) \propto \frac{1}{\ell^{2\gamma}}, \quad (3.37)$$

dando lugar a dos casos diametralmente opuestos:

- (i) $2\gamma < 1$. En ese caso $\Rightarrow \gamma < 1/2 \Rightarrow 2 - 2H < 1/2 \Rightarrow H > 3/4$. Las correlaciones de la serie módulo decaen más lentamente que $1/\ell$ dando lugar a correlaciones de largo alcance en la serie módulo.
- (ii) $2\gamma > 1$, por tanto, $H < 3/4$. En este caso, aunque las correlaciones siguen siendo positivas (habíamos visto en Eq. 3.35 que la serie módulo de un fGn nunca puede tener correlaciones negativas) decaen de forma más rápida que $1/\ell$ alcanzando rápidamente valores propios del ruido blanco (ver Fig. 3.2). De hecho, los métodos indirectos utilizados para el cálculo del exponente de correlación basados en el estudio de fluctuaciones de la serie temporal, e.g. el FA y el DFA, no son capaces de detectar correlaciones en la serie módulo cuando $H < 3/4$ (Carpena et al., 2017). En consecuencia, el DFA otorga resultados espurios en la Fig. 2.3 y en (Gómez-Extremera et al., 2016; Kalisky et al., 2005) para la serie módulo en la región $\alpha_{\text{in}} < 3/4$ (recordamos que $\alpha = H$ en la región $0 < H \leq 1$) ya que no es capaz de detectar las correlaciones que decaen más rápido que $1/\ell$ y asigna un valor $\alpha_{\text{mag}} = 0.5$ cuando en realidad la serie módulo sí posee correlaciones positivas, pero de corto alcance. El valor $\alpha_{\text{mag}} = 0.5$ obtenido a través del DFA en la región $0 < \alpha_{\text{in}} < 3/4$ únicamente es verdadero para $\alpha_{\text{in}} = 0.5$.

3.2. Autocorrelación de la serie cuadrática

Se ha probado numéricamente que las propiedades de scaling de la serie cuadrática $\{x_i^2\}$ coinciden con las de la serie módulo $\{|x_i|\}$ (Kalisky et al., 2005) (en esta sección lo probamos de manera analítica). Debido a esto, por simplicidad, en muchos casos se estudia la función de autocorrelación de la serie cuadrática en detrimento de la del módulo. La función de autocorrelación de la serie cuadrática viene dada por:

$$C_{x^2}(\ell) = \frac{\langle x_i^2 \cdot x_{i+\ell}^2 \rangle - \langle x_i^2 \rangle \langle x_{i+\ell}^2 \rangle}{\sigma_{x^2}^2} \quad (3.38)$$

A continuación, y de manera análoga a como hicimos en la sección anterior (Sec. 3.1) para el módulo, calculamos la relación analítica entre la función de autocorrelación de la

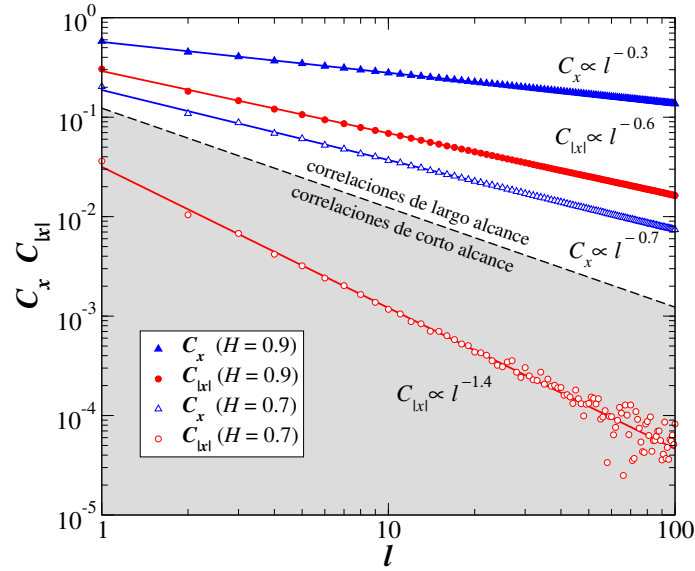


Figura 3.2: Función de autocorrelación C_x y función de autocorrelación de la serie módulo $C_{|x|}$ para dos fGns (uno con $H > 0.75$ y otro con $H < 0.75$) generados a través del método de filtrado de Fourier y con un tamaño de serie $N = 2^{24} \sim 1.6 \cdot 10^7$. Para evitar fluctuaciones estadísticas se promedia en cada caso sobre 100 experimentos. La línea discontinua corresponde a la frontera entre la zona de correlaciones de corto y largo alcance. En el caso de $H = 0.9$ vemos que la función de autocorrelación, C_x , y la de su módulo, $C_{|x|}$ presentan correlaciones de largo alcance ya que $\gamma < 1$ y, por tanto, la función de autocorrelación decae más lentamente que $1/\ell$. En cambio, para $H = 0.7$ la función de autocorrelación, C_x muestra correlaciones de largo alcance (de nuevo $\gamma < 1$) mientras que la de su módulo, $C_{|x|}$, cae en el régimen de correlaciones de corto alcance ya que $\gamma > 1$ por lo que la función de autocorrelación decae más rápido que $1/\ell$. Además, en este ejemplo en particular debido a la rápida caída de $C_{|x|}$ y al bajo valor que toma $C_{|x|}$ para $\ell = 1$, el nivel del ruido se alcanza para distancias relativamente cortas ($\ell < 100$) incluso para series temporales tan largas (se considera que una función de autocorrelación entra en la zona de ruido cuando $C(\ell) < 1/\sqrt{N}$). En ambos casos, $H = 0.7$ y $H = 0.9$, el exponente de caída de la función de autocorrelación del módulo toma el valor doble del exponente de la función de autocorrelación de la serie original, de acuerdo con la Eq. 3.37.

serie cuadrática, C_x^2 , y la función de autocorrelación de la serie original, C_x , para ruidos fraccionarios Gaussianos.

La covarianza de la serie cuadrática $K_{sq}(\ell)$ para dos variables aleatorias X, Y normalmente distribuidas vendría dada por la expresión:

$$K_{sq}(\ell) = \langle x^2 \cdot y^2 \rangle - \langle x^2 \rangle \langle y^2 \rangle, \quad (3.39)$$

y expresada en forma integral:

$$K_{sq}(\ell) = \int_{-\infty}^{\infty} x^2 dx \int_{-\infty}^{\infty} y^2 dy \cdot \rho(x, y) - 1 \quad (3.40)$$

siendo de nuevo $\rho(x, y)$ la distribución Gaussiana bivariada y $\langle x^2 \rangle = \langle y^2 \rangle = 1$. Teniendo en cuenta la Eq. 3.1 y denotando $A_k = \sqrt{1 - k_x^2(\ell)}$ podemos expresar la autocovarianza de la serie cuadrática mediante la expresión:

$$K_{sq}(\ell) = \frac{1}{2\pi A_k} \int_{-\infty}^{\infty} x^2 dx \int_{-\infty}^{\infty} y^2 dy \cdot \exp\left(-\frac{x^2 + y^2 - 2k_x(\ell)xy}{2A_k^2}\right) - 1. \quad (3.41)$$

Realizando el cambio de variable:

$$\begin{aligned} \xi &= \frac{x}{A_k}; \quad x = \xi \cdot A_k; \quad dx = A_k d\xi \\ \varphi &= \frac{y}{A_k}; \quad y = \varphi \cdot A_k; \quad dy = A_k d\varphi \end{aligned} \quad (3.42)$$

expresamos la covarianza de la serie cuadrática como:

$$K_{sq}(\ell) = \frac{A_k^5}{2\pi} \int_{-\infty}^{\infty} \xi^2 d\xi \exp\left(-\frac{\xi^2}{2}\right) \int_{-\infty}^{\infty} \varphi^2 d\varphi \exp\left(-\frac{\varphi^2}{2} + k_x(\ell)\xi\varphi\right) d\varphi \quad (3.43)$$

Llamamos I_1 a la segunda integral que aparece en (3.43), es decir:

$$I_1 = \int_{-\infty}^{\infty} \varphi^2 d\varphi \exp\left(-\frac{\varphi^2}{2} + k_x(\ell)\xi\varphi\right) d\varphi, \quad (3.44)$$

y la resolvemos utilizando la siguiente identidad (Gradshteyn and Ryzhik, 2014):

$$\int_{-\infty}^{\infty} x^2 \exp(-\mu x^2 + 2\nu x) dx = \frac{1}{2\mu} \sqrt{\frac{\pi}{\mu}} \left(1 + 2\frac{\nu^2}{\mu}\right) \exp\left(\frac{\nu^2}{\mu}\right). \quad (3.45)$$

En nuestro caso $\mu = 1/2$, y $2\nu = k_x(\ell)\xi \Rightarrow \nu = k_x(\ell)\xi/2$, por lo que al sustituir obtenemos

un valor para la integral I_1 :

$$I_1 = \sqrt{2\pi} \left(1 + k_x^2(\ell)\xi^2\right) \exp\left(\frac{k_x^2(\ell)\xi^2}{2}\right), \quad (3.46)$$

y por tanto, la covarianza de la serie cuadrática puede ser expresada en términos de una integral en una única variable ξ :

$$K_{sq}(\ell) = \frac{A_k^5}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \xi^2 \exp\left(-\frac{A_k^2}{2}\xi^2\right) d\xi + \frac{A_k^5}{\sqrt{2\pi}} k_x^2(\ell) \int_{-\infty}^{\infty} \xi^4 \exp\left(-\frac{A_k^2}{2}\xi^2\right) d\xi - 1. \quad (3.47)$$

Deshacemos el cambio de variable realizado en (3.42), i.e. $\xi = x/A_k \Rightarrow d\xi = dx/A_k$ obteniendo:

$$\begin{aligned} K_{sq}(\ell) &= \frac{A_k^2}{\sqrt{2\pi}} \underbrace{\int_{-\infty}^{\infty} x^2 \exp\left(-\frac{x^2}{2}\right) dx}_{I_2} + \\ &\quad \frac{k_x^2(\ell)}{\sqrt{2\pi}} \underbrace{\int_{-\infty}^{\infty} x^4 \exp\left(-\frac{x^2}{2}\right) dx}_{I_3} - 1, \end{aligned} \quad (3.48)$$

y resolvemos I_2 utilizando la identidad mostrada en (3.45) usando $\mu = 1/2$ y $\nu = 0$ obteniendo $I_2 = \sqrt{2\pi}$. Para la resolución de la integral I_3 utilizamos la siguiente identidad, encontrada en (Gradshteyn and Ryzhik, 2014):

$$\int_{-\infty}^{\infty} x^n \exp(-px^2 + 2qx) dx = \frac{1}{2^{n-1}p} \sqrt{\frac{\pi}{p}} \frac{d^{n-1}}{dq^{n-1}} \left(q \exp\left(\frac{q^2}{p}\right) \right), \quad p > 0, \quad (3.49)$$

siendo $n = 2$, $q = 0$, $p = 1/2$ y, por tanto:

$$\frac{d^{n-1}}{dq^{n-1}} \left(q e^{q^2/p} \right) = \frac{2}{p} \left[e^{q^2/p} + \frac{2q^2}{p} e^{q^2/p} \cdot (3 + q^2) + 2q^3 e^{q^2/p} \right] = 12, \quad (3.50)$$

por lo que finalmente la integral I_3 toma el valor:

$$I_3 = \frac{1}{2^3 \cdot \frac{1}{2}} \sqrt{2\pi} \cdot 12 = 3\sqrt{2\pi}, \quad (3.51)$$

reemplazamos en (3.48):

$$K_{sq}(\ell) = \frac{A_k^2}{\sqrt{2\pi}} \sqrt{2\pi} + \frac{k_x^2(\ell)}{\sqrt{2\pi}} \cdot 3\sqrt{2\pi} - 1 = 1 - k_x^2(\ell) + 3k_x^2(\ell) - 1, \quad (3.52)$$

y finalmente:

$$K_{sq}(\ell) = 2k_x^2(\ell) \quad (3.53)$$

Ahora bien, teniendo en cuenta que para la serie cuadrática $\{x_i^2\}$ la desviación estándar es $\sigma_{x^2} = \sqrt{2}$, su función de autocorrelación vendrá dada por:

$$C_{x^2}(\ell) = \frac{K_{sq}(\ell)}{\sigma_{x^2}^2} = \frac{2k_x^2(\ell)}{\sqrt{2}\sqrt{2}} = k_x^2(\ell). \quad (3.54)$$

Teniendo en cuenta que la serie de partida $\{x_i\}$ es un ruido lineal Gaussiano con media nula y desviación estándar unidad, la función de autocorrelación de la serie cuadrática toma la forma:

$$C_{x^2}(\ell) = C_x^2(\ell) \quad (3.55)$$

Este resultado implica que la autocorrelación de la serie cuadrática, al igual que la del módulo, nunca puede tomar valores negativos (ver Fig. 3.3). Además, este resultado también explica el motivo por el que la función de autocorrelación del módulo y la serie cuadrática escalan con el mismo exponente (asintóticamente), i.e. tienen las mismas propiedades de scaling, (Kalisky et al., 2005) ya que para grandes valores del retardo ℓ la autocorrelación de la serie original toma valores muy pequeños $C_x \ll 1$, propiciando que la autocorrelación del módulo y la serie cuadrática tomen valores prácticamente idénticos. Otra forma de corroborar esto consiste en considerar la aproximación (3.36). Es evidente que si $C_{|x|}$ y C_{x^2} siguen una ley de potencias, ambas escalarán con el mismo exponente. Sin embargo, dejando de lado el exponente de scaling claramente influenciado por la cola de la función de autocorrelación (grandes valores del retardo ℓ), el uso de la función de autocorrelación de la serie cuadrática en detrimento del de la serie módulo, puede inducirnos a errores a la hora de estimar valores de autocorrelación del módulo a distancias cortas. En la Fig. 3.3 podemos ver que el error relativo entre ambas funciones de autocorrelación puede alcanzar valores hasta de un 20 % para valores $C_x \simeq \pm 0.5$. Este hecho ensalza aún más el resultado obtenido en la Eq. 3.36, ya que usando esta aproximación cometeríamos un error relativo aproximado del 2 % para valores $C_x \simeq \pm 0.5$. Por tanto, concluimos que a pesar de que la serie cuadrática y la serie módulo muestran propiedades de scaling similares, no es conveniente sustituir la serie módulo por la serie cuadrática para el estudio de la autocorrelación a corta escala del módulo. Por el contrario, la aproximación encontrada en (3.36) funciona con gran precisión hasta valores cercanos a $C_x \sim \pm 0.75$.

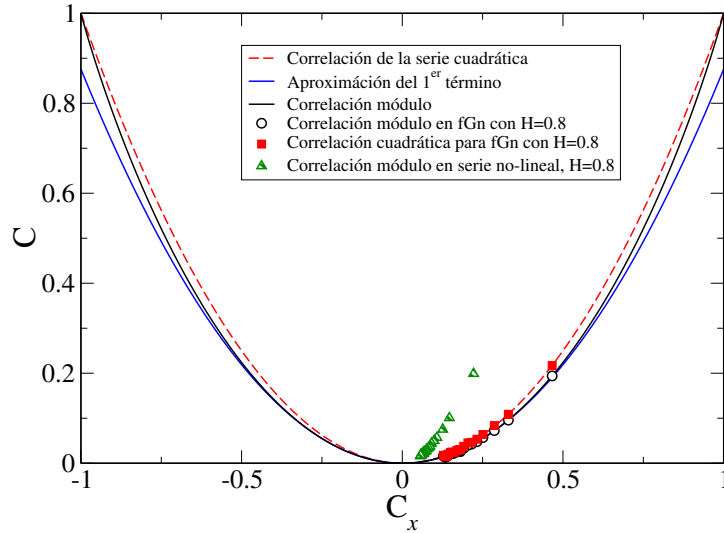


Figura 3.3: Autocorrelación de la serie módulo $C_{|x|}$ y la serie cuadrática C_{x^2} en función de la autocorrelación de la serie original C_x . La línea continua negra corresponde a la autocorrelación de la serie módulo $C_{|x|}$ obtenida en la Eq. 3.35 para ruidos fraccionarios Gaussianos mientras que la línea continua azul corresponde al primer término de su aproximación mediante desarrollo en serie de potencias (3.36). La línea discontinua roja representa la autocorrelación de la serie cuadrática C_{x^2} obtenida en (3.55). Los símbolos blancos corresponden a la autocorrelación del módulo de un fGN $\{x_i\}$ generado a través del método de Filtrado de Fourier con un exponente de Hurst $H = 0.8$ y un tamaño de secuencia $N = 2^{18}$, mientras que los símbolos rojos representan la autocorrelación de la serie cuadrática $\{x_i^2\}$. Por último, los triángulos verdes representan la autocorrelación del módulo para un ruido no lineal que consta del mismo exponente de correlación lineal que los ruidos generados anteriormente ($H=0.8$). Se aprecia una clara desviación sobre la curva teórica (Eq. 3.35) para ruidos lineales Gaussianos.

3.3. Autocorrelación de la serie signo

Como ya vimos en el capítulo anterior, la serie módulo y la serie signo son una gran fuente de información acerca de la señal original. En concreto, dada una serie original $\{x_i\}$ el módulo indica cómo de grandes son los cambios mientras que el signo se encarga de facilitar su dirección. La serie signo es de gran importancia también a la hora de estudiar los tiempos de paso en procesos correlacionados (Carretero-Campos et al., 2012), en búsqueda de palabras clave (Carpena et al., 2009; Ortuno et al., 2002), y en el estudio de actividad sísmica (Corral, 2006). Partiendo de un fGn $\{x_i\}$ Apostolov et al. (Apostolov et al., 2008) han obtenido una relación analítica para la autocorrelación de la serie signo en función de la autocorrelación de la serie original para series Gaussianas:

$$C_x = \sin\left(\frac{\pi}{2}C_s\right) \quad (3.56)$$

donde C_s es la autocorrelación de la serie signo, y por tanto:

$$C_s = \frac{2}{\pi} \arcsin(C_x). \quad (3.57)$$

A continuación obtenemos la relación analítica obtenida por Apostolov a través de dos métodos diferentes:

Integración

La autocorrelación de la serie signo de una serie Gaussiana con media nula y desviación estándar unidad entre dos variables ($\text{sgn}(x), \text{sgn}(y)$) separadas temporalmente por ℓ puede definirse como:

$$C_s = \frac{\langle \text{sgn}(x) \cdot \text{sgn}(y) \rangle - \langle \text{sgn}(x) \rangle \langle \text{sgn}(y) \rangle}{\sigma_{\text{sgn}(x)} \sigma_{\text{sgn}(y)}} = \langle \text{sgn}(x) \cdot \text{sgn}(y) \rangle, \quad (3.58)$$

por lo que la expresión en forma de integral de C_s será:

$$C_s = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{sgn}(x) \text{sgn}(y) h(x, y, k) dx dy \quad (3.59)$$

donde $h(x, y, k)$ es la distribución Gaussiana bivariada, por lo que:

$$C_s = \frac{1}{2\pi A_k} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{sgn}(x) \text{sgn}(y) \exp\left(-\frac{x^2 + y^2 - 2kxy}{2A_k^2}\right) dx dy, \quad (3.60)$$

siendo $A_k = \sqrt{1 - k^2}$. Para resolver la integral en primer lugar consideramos la integral en la variable y , es decir:

$$I = \int_{-\infty}^{\infty} \operatorname{sgn}(y) \exp\left(-\frac{x^2 + y^2 - 2kxy}{2A_k^2}\right) dy \quad (3.61)$$

$$= \int_{-\infty}^0 -1 \cdot \exp\left(-\frac{x^2 + y^2 - 2kxy}{2A_k^2}\right) dy + \int_0^{\infty} \exp\left(-\frac{x^2 + y^2 - 2kxy}{2A_k^2}\right) dy. \quad (3.62)$$

Expresamos (3.62) de forma que aparezca un cuadrado perfecto en las exponenciales:

$$I = \exp\left(\frac{-x^2 + k^2x^2}{2A_k^2}\right) \left[\underbrace{\int_0^{\infty} \exp\left(-\left(\frac{y - kx}{\sqrt{2}A_k}\right)^2\right) dy - \int_{-\infty}^0 \exp\left(-\left(\frac{y - kx}{\sqrt{2}A_k}\right)^2\right) dy}_{I_1} \right] \quad (3.63)$$

A continuación, realizamos el siguiente cambio de variable para poder expresar I_1 en función de la función error:

$$t = \frac{y - kx}{\sqrt{2}A_k}; \quad dt = \frac{1}{\sqrt{2}A_k} dy \Rightarrow dy = \sqrt{2}A_k dt, \quad (3.64)$$

obteniendo para I_1 :

$$I_1 = \sqrt{2}A_k \left[\int_{-kx/(\sqrt{2}A_k)}^{\infty} \exp(-t^2) dt - \int_{-\infty}^{-kx/(\sqrt{2}A_k)} \exp(-t^2) dt \right], \quad (3.65)$$

y teniendo en cuenta que la función error, $\operatorname{erf}(x)$ y la función error complementario, $\operatorname{erfc}(x)$, son funciones impares que vienen dadas por:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt, \quad \operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt, \quad (3.66)$$

estaríamos en disposición de resolver I_1 :

$$I_1 = \sqrt{2}A_k \left[\cancel{\frac{-\sqrt{\pi}}{2} \operatorname{erfc}\left(\frac{kx}{\sqrt{2}A_k}\right)} + \sqrt{\pi} \cdot \operatorname{erf}\left(\frac{kx}{\sqrt{2}A_k}\right) + \cancel{\frac{\sqrt{\pi}}{2} \operatorname{erfc}\left(\frac{kx}{\sqrt{2}A_k}\right)} \right], \quad (3.67)$$

y reemplazando finalmente I_1 en (3.63):

$$I = \exp\left(\frac{-x^2 + k^2x^2}{2A_k^2}\right) \sqrt{2\pi}A_k \cdot \operatorname{erf}\left(\frac{kx}{\sqrt{2}A_k}\right). \quad (3.68)$$

Para completar la integración, debemos realizar la integral en la variable x en (3.60):

$$\begin{aligned} C_s &= \frac{1}{2\pi A_k} \sqrt{2\pi} \int_{-\infty}^{\infty} \text{sgn}(x) \exp\left(\frac{-x^2 + k^2 x^2}{2A_k^2}\right) \text{erf}\left(\frac{kx}{\sqrt{2}A_k}\right) dx \\ &= \frac{1}{\sqrt{2\pi}} \left[\int_0^{\infty} \exp\left(\frac{-x^2 + k^2 x^2}{2A_k^2}\right) \text{erf}\left(\frac{kx}{\sqrt{2}A_k}\right) dx - \int_{-\infty}^0 \exp\left(\frac{-x^2 + k^2 x^2}{2A_k^2}\right) \text{erf}\left(\frac{kx}{\sqrt{2}A_k}\right) dx \right] \end{aligned} \quad (3.69)$$

La función resultante en ambos integrandos es una función impar, debido a que se trata del producto entre una función impar (función error) y una función par (exponencial). Por tanto, la correlación del signo, C_s se puede expresar como:

$$C_s = \sqrt{\frac{2}{\pi}} \int_0^{\infty} \exp\left(\frac{-x^2 + k^2 x^2}{2A_k^2}\right) \text{erf}\left(\frac{kx}{\sqrt{2}A_k}\right) dx. \quad (3.70)$$

Para su resolución, usamos la siguiente identidad encontrada en (Ng and Geller, 1969):

$$\int_0^{\infty} \text{erf}(ax) e^{-b^2 x^2} dx = \frac{\sqrt{\pi}}{2b} - \frac{1}{b\sqrt{\pi}} \arctan\left(\frac{b}{a}\right), \quad (3.71)$$

siendo en nuestro caso:

$$b = \pm \frac{1}{\sqrt{2}}, \quad a = \frac{k}{\sqrt{2}A_k}. \quad (3.72)$$

Para que la correlación del signo tome valores en el intervalo $[-1, 1]$ (única solución aceptable) debemos tomar $b = 1/\sqrt{2}$ si $k \geq 0$ y $b = -1/\sqrt{2}$ si $k < 0$, dando lugar a dos posibles casos:

$$C_s = \begin{cases} \langle \text{sgn}(x) \cdot \text{sgn}(y) \rangle = 1 - \frac{2}{\pi} \arctan\left(\frac{A_k}{k}\right) & \text{si } k \geq 0 \\ \langle \text{sgn}(x) \cdot \text{sgn}(y) \rangle = -1 - \frac{2}{\pi} \arctan\left(\frac{A_k}{k}\right) & \text{si } k < 0 \end{cases}$$

Ahora bien, con objeto de llegar a la misma expresión obtenida por Apostolov (Eq. 3.57), busquemos una expresión que relacione las funciones inversas del seno y la tangente. Consideremos un ángulo α de forma que $\tan(\alpha) = t \Leftrightarrow \tan^{-1}(t) = \alpha$. Utilizando la conocida relación trigonométrica $\tan^2(\alpha) + 1 = 1/\cos^2(\alpha)$ podemos obtener una relación entre arcoseno y arcotangente:

$$\arcsen\left(\pm \sqrt{\frac{t^2}{1+t^2}}\right) = \arctan(t), \quad (3.73)$$

siendo en nuestro caso $t = A_k/k$, por lo que:

$$\arctan\left(\frac{A_k}{k}\right) = \arcsen\left(\pm\sqrt{1-k^2}\right), \quad (3.74)$$

y teniendo en cuenta que el arcoseno y el arcotangente son funciones impares podemos expresar la autocorrelación en la serie signo como:

$$C_s = \begin{cases} \langle \text{sgn}(x) \cdot \text{sgn}(y) \rangle = 1 - \frac{2}{\pi} \arcsen(\sqrt{1-k^2}) & \text{si } k \geq 0 \\ \langle \text{sgn}(x) \cdot \text{sgn}(y) \rangle = -1 - \frac{2}{\pi} \arcsen(-\sqrt{1-k^2}) & \text{si } k < 0 \end{cases}$$

Finalmente, para obtener la Eq. 3.57 hacemos una manipulación trigonométrica:

(i) $k \geq 0$. Consideramos $\text{sen}(z) = k$ por lo que $\sqrt{1-k^2} = \cos(z)$, y por tanto:

$$C_s = 1 - \frac{2}{\pi} \arcsen(\sqrt{1-k^2}) = 1 - \frac{2}{\pi} \arcsen(\cos(z)). \quad (3.75)$$

Teniendo en cuenta que $k \geq 0 \Rightarrow 0 \leq z \leq \pi/2$ y podemos establecer la relación:

$$\cos(z) = \text{sen}\left(\frac{\pi}{2} - z\right). \quad (3.76)$$

Reemplazando (3.76) en (3.75) obtenemos:

$$C_s = 1 - \frac{2}{\pi} \arcsen\left(\text{sen}\left(\frac{\pi}{2} - z\right)\right) = \frac{2}{\pi} z. \quad (3.77)$$

Deshaciendo el cambio de variable ($z = \arcsen(k)$) la expresión de la autocorrelación del signo viene dada por:

$$C_s = \frac{2}{\pi} \arcsen(k) \quad k \geq 0 \quad (3.78)$$

(ii) $k < 0$. De nuevo, introducimos el cambio de variable $\text{sen}(z) = k \Rightarrow \sqrt{1-k^2} = \cos(z)$, teniendo en este caso la relación:

$$C_s = -1 - \frac{2}{\pi} \arcsen(-\sqrt{1-k^2}) = -1 - \frac{2}{\pi} \arcsen(-\cos(z)). \quad (3.79)$$

Considerando $k < 0 \Rightarrow -\frac{\pi}{2} \leq z < 0$ y podemos establecer la relación:

$$\cos(z) = \text{sen}\left(\frac{\pi}{2} + z\right). \quad (3.80)$$

Reemplazando (3.80) en (3.79) y haciendo uso del carácter impar de la función

arcoseno:

$$C_s = -1 + \frac{2}{\pi} \arcsen \left(\sin \left(\frac{\pi}{2} + z \right) \right) = \frac{2}{\pi} z. \quad (3.81)$$

De nuevo, deshaciendo el cambio de variable $\sin(z) = k \Leftrightarrow z = \arcsen(k)$ obtenemos

$$C_s = \frac{2}{\pi} \arcsen(k) \quad k < 0 \quad (3.82)$$

Por último, generalizando (3.78) y (3.82), y teniendo en cuenta que nuestras series de partida tienen media nula y desviación estándar unidad (autocovarianza k y autocorrelación C_x coinciden) obtenemos la relación de Apostolov (3.57):

$$C_s = \frac{2}{\pi} \arcsen(C_x) \quad \forall C_x \in [-1, 1] \quad (3.83)$$

Transformación afín

Mediante este procedimiento vamos a probar que la relación de Apostolov es válida para todas las series con distribución de valores simétrica (no necesariamente ha de tener una distribución de valores Gaussiana). Dadas dos variables aleatorias x_1 y x_2 , definimos la correlación de su signo:

$$\begin{aligned} C_s &= \frac{\langle \text{sgn}(x_1) \cdot \text{sgn}(x_2) \rangle - \langle \text{sgn}(x_1) \rangle \langle \text{sgn}(x_2) \rangle}{\sigma_{\text{sgn}(x_1)} \sigma_{\text{sgn}(x_2)}} = \langle \text{sgn}(x_1) \cdot \text{sgn}(x_2) \rangle \\ &= 1 \cdot \text{Prob}[\text{sgn}(x_1) = \text{sgn}(x_2)] - 1 \cdot \text{Prob}[\text{sgn}(x_1) \neq \text{sgn}(x_2)] \\ &= 1 \cdot \text{Prob}(x_1 > 0, x_2 > 0) + 1 \cdot \text{Prob}(x_1 < 0, x_2 < 0) \\ &\quad - 1 \cdot \text{Prob}(x_1 > 0, x_2 < 0) - 1 \cdot \text{Prob}(x_1 < 0, x_2 > 0) \\ &= 2 \cdot \text{Prob}(x_1 > 0, x_2 > 0) - [1 - 2 \cdot \text{Prob}(x_1 > 0, x_2 > 0)]. \end{aligned} \quad (3.84)$$

Denotando $P = \text{Prob}(x_1 > 0, x_2 > 0)$, la autocorrelación de la serie signo vendrá dada por $C_s = 2P - (1 - 2P) = 4P - 1$. Ahora bien, si x_1 y x_2 siguen una distribución elíptica¹ (McNeil et al., 2015) la matriz de covarianza Σ entre las dos variables x_1 y x_2 puede expresarse como:

$$\Sigma = \begin{pmatrix} 1 & k \\ k & 1 \end{pmatrix} \quad (3.85)$$

¹Las variables correlacionadas x_1 y x_2 siguen una distribución elíptica cuando sus distribuciones marginales son simétricas

donde k es la covarianza entre x_1 y x_2 . Si una variable X sigue una distribución elíptica se puede establecer una transformación afín:

$$X = \mu + A \cdot Z \quad (3.86)$$

donde A es una matriz 2×2 , μ es la media (nula en nuestro caso) y Z es una variable aleatoria que sigue distribución esférica (McNeil et al., 2015) y por tanto, cuya matriz de covarianza es la matriz identidad². La matriz de covarianza de la variable X viene dada por $\mathcal{N}(\mu, \Sigma) = \mathcal{N}(\mu, AA^t)$. A partir de la factorización de Cholesky de Σ podemos obtener los valores de la matriz A :

$$\begin{pmatrix} 1 & k \\ k & 1 \end{pmatrix} = \begin{pmatrix} a & 0 \\ c & d \end{pmatrix} \begin{pmatrix} a & c \\ 0 & d \end{pmatrix} = \begin{pmatrix} a^2 & ac \\ ca & c^2 + d^2 \end{pmatrix}, \quad (3.87)$$

y tras resolver el sistema de ecuaciones obtenemos:

$$A = \begin{pmatrix} 1 & 0 \\ k & \sqrt{1-k^2} \end{pmatrix}, \quad (3.88)$$

y reemplazando en Eq. 3.86:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ k & \sqrt{1-k^2} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \quad (3.89)$$

dando lugar al siguiente sistema de ecuaciones:

$$\left. \begin{aligned} z_1 &= x_1 \\ kz_1 + z_2\sqrt{1-k^2} &= x_2 \end{aligned} \right\} \quad (3.90)$$

Dado nuestro interés en el cálculo de P , nuestro sistema de ecuaciones se transforma en:

$$\left. \begin{aligned} z_1 &> 0 \\ kz_1 + z_2\sqrt{1-k^2} &> 0 \end{aligned} \right\} \quad (3.91)$$

Haciendo un doble cambio a coordenadas polares: $\{z_1 = R\cos(\phi), z_2 = R\sin(\phi), R > 0\}$ y $\{k = \sin(\theta), \sqrt{1-k^2} = \cos(\theta)\}$ obtenemos:

$$\left. \begin{aligned} \cos(\phi) &> 0 \\ \sin(\theta)\cos(\phi) + \sqrt{1-k^2}\sin(\theta) &> 0 \end{aligned} \right\} \quad (3.92)$$

²Un ejemplo sería la $\mathcal{N}(0, 1)$

por lo que la solución es $\phi > 0$, $\phi > -\theta$, y por tanto, podemos obtener P mediante el cociente entre el área que representa esta región y el área total del círculo, es decir:

$$P = \frac{\pi/2 + \theta}{2\pi}. \quad (3.93)$$

Finalmente calculamos la correlación del signo C_s :

$$C_s = 4P - 1 = \frac{2\pi + 4\theta}{2\pi} - 1 = \frac{2}{\pi}\theta = \frac{2}{\pi}\arcsen(k). \quad (3.94)$$

En el caso de series con media nula y varianza unidad la autocovarianza k coincide con la autocorrelación C_x y obtenemos la Eq. 3.57:

$$C_s = \frac{2}{\pi}\arcsen(C_x) \quad (3.95)$$

Una vez probada la relación de Apostolov de dos formas diferentes, además de ver que no solo es válida para series con distribución de valores Gaussianos sino que lo es para cualquier serie con distribución de valores simétrica, estudiamos sus propiedades. Realizando un desarrollo en serie de potencias para pequeños valores de ℓ en 3.57, obtenemos:

$$C_s = \frac{2}{\pi}C_x + \mathcal{O}(C_x^3). \quad (3.96)$$

En la Fig. 3.4 podemos ver la dependencia funcional tanto de la autocorrelación del signo como la del primer término de su desarrollo en serie de potencias en función de la autocorrelación de la serie original. Esta aproximación (Eq. 3.96) nos indica que la autocorrelación de la serie signo es directamente proporcional a la de la señal original. Además, en el caso de ruidos fraccionarios Gaussianos, C_x y C_s decaen en forma de ley de potencias (ver Sec. 2.1) por lo que ambas deben mostrar el mismo exponente de scaling. A raíz de este resultado, reemplazamos (1.12) en (3.96) obteniendo:

$$C_s(\ell) \propto \frac{1}{\ell^\gamma}, \quad (3.97)$$

dando lugar a dos casos de interés (ver Fig. 3.5):

- (i) $\gamma < 1$. En ese caso, teniendo en cuenta (1.14), $\alpha_s > 0.5$. Las correlaciones decaen más lento que $1/\ell$ por lo que tendremos correlaciones positivas de largo alcance. Además, el exponente de correlación α de la serie original coincide con el de la serie signo α_s .
- (ii) $\gamma > 1$. Teniendo de nuevo en cuenta (1.14) para este caso tendríamos que $\alpha_s < 0.5$,

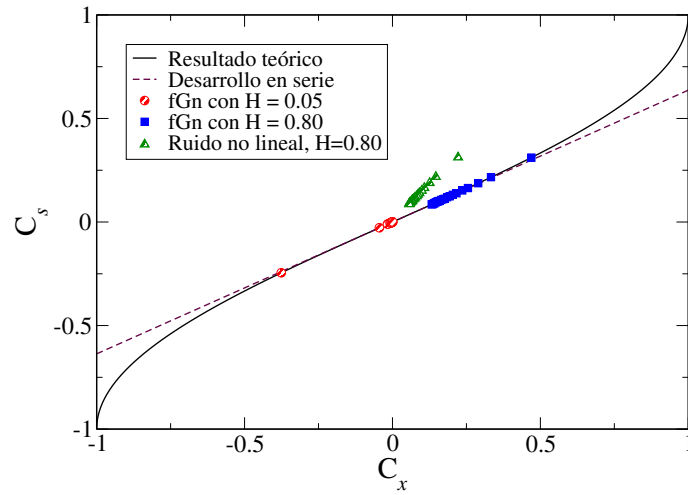


Figura 3.4: Autocorrelación de la serie signo C_s en función de la autocorrelación de la serie original C_x . La línea continua representa la autocorrelación de la serie signo C_s obtenida en Eq. 3.57 mientras que la línea discontinua corresponde al primer término de su aproximación mediante un desarrollo en serie de potencias (3.96). Dicha aproximación funciona correctamente siempre y cuando el valor absoluto de la correlación de la serie original no sea excesivamente alto. De hecho, para valores $C_x \approx \pm 0.4$ el error relativo no supera el 3%. Los círculos corresponden a un fGn con $H=0.05$, los cuadrados a un fGn con $H=0.80$ y los triángulos a un ruido no-lineal generado mediante composición de módulos y signos independientes, con correlaciones lineales $H = 0.8$, donde se puede apreciar una clara desviación sobre la curva de la Eq. 3.57. Todos los ruidos tienen un tamaño de $N = 2^{20} \simeq 10^6$.

lo que implica que las correlaciones en el signo decaen más rápido que $1/\ell$ por lo que serán correlaciones negativas y de corto alcance. Considerando la Eq. 3.96, el exponente de correlación de la serie signo debe ser en todo momento igual al de la serie original. Sin embargo, en la Sec. 2.1 obtuvimos mediante el uso del DFA que los ruidos fraccionarios Gaussianos anticorrelacionados ($\alpha < 0.5$) poseen signos aleatorios ($\alpha_s=0.5$). De nuevo estamos ante un resultado espurio del DFA (Carpena et al., 2017) que no es capaz de detectar las correlaciones de corto alcance existentes en las series signo cuando $\alpha < 0.5$, ya que teniendo en cuenta la Eq. 3.96 el exponente de la serie signo de un fGn debe coincidir siempre con el de su serie original.

Finalmente, reemplazando Eq. 3.56 en Eq. 3.35 podemos obtener la dependencia analítica de la autocorrelación del módulo en función de la del signo para fGns:

$$C_{|x|} = \frac{2}{\pi - 2} \left[\frac{\pi}{2} C_s \sin\left(\frac{\pi}{2} C_s\right) + \cos\left(\frac{\pi}{2} C_s\right) - 1 \right], \quad (3.98)$$

y su desarrollo en serie de potencias para pequeños valores de ℓ :

$$C_{|x|} = \frac{\pi^2}{4(\pi - 2)} C_s^2 + \mathcal{O}(C_s^4) \quad (3.99)$$

La desviación de (3.98) puede ser utilizada como un indicador de acoplamiento lineal entre el módulo y el signo y puede ser de gran ayuda a la hora de estudiar sistemas en los que el comportamiento del módulo y el signo juegan un papel esencial en la dinámica del sistema. En la Fig. 3.6 podemos ver la dependencia funcional de las Eqs. 3.98 y 3.99 junto a los resultados obtenidos para dos ruidos (uno presenta acoplamiento lineal entre módulo y signo mientras que el otro no). Se puede apreciar que para el caso del ruido con acoplamiento lineal los puntos siguen perfectamente la curva teórica (3.98), mientras que para el ruido sin acoplamiento lineal, a pesar de mostrar exactamente las mismas propiedades de scaling que el ruido con acoplamiento lineal, los puntos están totalmente desviados de la curva teórica para el acoplamiento lineal (Eq. 3.98).

3.4. Modelo no lineal de series

Hasta el momento, hemos estudiado únicamente la relación entre la autocorrelación de señal original, modulo y signo para ruidos lineales Gaussianos (Eqs. 3.35 y 3.57). Sin embargo, estas ecuaciones dejan de ser válidas cuando estudiamos ruidos Gaussianos no lineales, y la desviación sobre dichas ecuaciones puede ser utilizada como una medida de no-linealidad. Cabe destacar que esta nueva medida de no-linealidad no requiere la presencia alguna de scaling en la serie bajo estudio, lo que supone una gran ventaja frente a los métodos tradicionales utilizados para evaluar las correlaciones del módulo (técnicas de Análisis de Fluctuación y Multifractal DFA) en series naturales, donde la presencia de scaling no está siempre garantizada. En concreto, vamos a centrar nuestra atención en la Eq. 3.35, debido a que diversos estudios muestran que las propiedades no lineales y multifractales de una serie están asociadas a las propiedades de correlación de la serie módulo (Ashkenazy, 2001; Gómez-Extremera et al., 2016; Kalisky et al., 2005) (ya hemos visto que esto no es siempre así, ya que ruidos correlacionados pueden presentar correlaciones en el módulo, pero ajustarse perfectamente a la curva Eq. 3.35). Para mostrar el efecto de la no-linealidad en series temporales vamos a generar series con propiedades no-lineales mediante la composición de módulos y signos independientes (tanto el módulo como el signo proceden de series con distribución de valores Gaussiana)(ver Sec. 2.2) inicialmente propuesta por (Kalisky et al., 2005). Recordamos que mediante este método obteníamos una serie compuesta $x(i)$ a través del producto de la serie módulo $x_{\text{mod}}(i)$ y

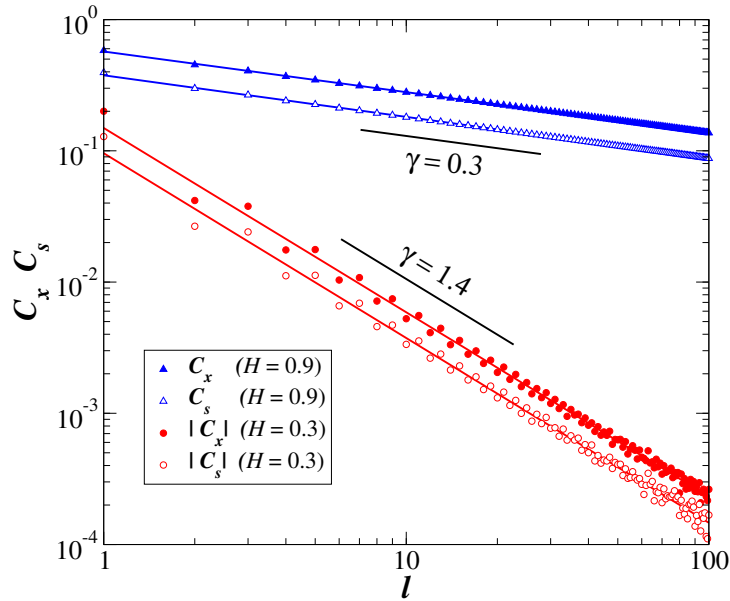


Figura 3.5: Función de autocorrelación de serie original C_x y serie signo C_s para dos ruidos fraccionarios Gaussianos con $\gamma = 0.3$ (correlacionado) y $\gamma = 1.4$ (anticorrelacionado). En este último caso ($\gamma = 1.4$) representamos $|C_x|$ y $|C_s|$ para permitir la representación en doble escala logarimica. Ambos fGns tienen un tamaño aproximado de 10^7 y están generados a través del Método de Filtrado de Fourier. Además, para evitar fluctuaciones estadísticas, las funciones de autocorrelación son fruto del promedio sobre 100 experimentos. En ambos casos, la función de autocorrelación de serie original y de serie signo decaen con el mismo exponente (son paralelas en doble escala logarítmica) de acuerdo con la Eq. 3.96. En el caso de la serie anticorrelacionada, $\gamma > 1$, la función de autocorrelación alcanza la franja del ruido para valores $\ell < 100$, ya que dicha franja se alcanza cuando $C_x < 1/\sqrt{N} \approx 3 \cdot 10^{-4}$.

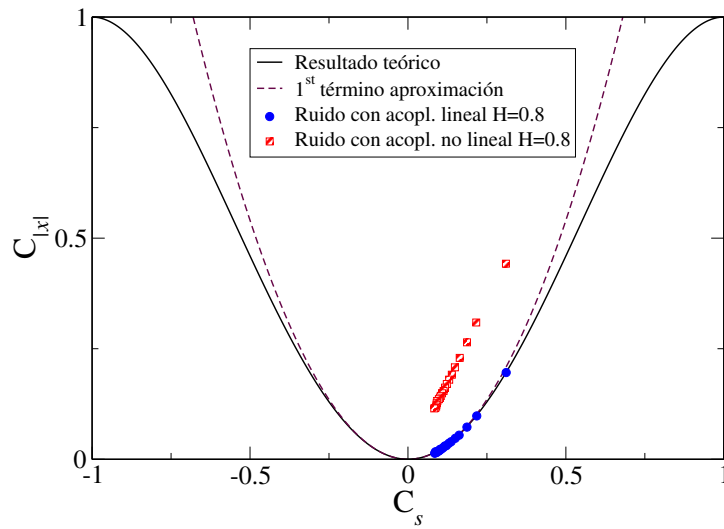


Figura 3.6: La línea sólida representa la autocorrelación de la serie módulo, $C_{|x|}$, en función de la autocorrelación de la serie signo C_s en ruidos lineales (Eq. 3.98). La línea discontinua representa el primer término no nulo del desarrollo en serie de potencias de (3.98) para valores ℓ pequeños. Los círculos azules representan los resultados experimentales obtenidos para un fGn con exponente de Hurst $H = 0.8$ y tamaño $N = 2^{20}$, mientras que los cuadrados rojos representan un ruido con $H = 0.8$ y acoplamiento entre módulo y signo no lineal generado mediante la composición de módulos y signos independientes (la distribución de valores sigue siendo Gaussiana). Para el caso del ruido lineal, los puntos experimentales se ajustan perfectamente a la curva teórica, mientras que para el ruido con acoplamiento no lineal los resultados experimentales están totalmente alejados de la curva predicha por Eq. 3.98 para el acoplamiento lineal.

la serie signo $x_s(i)$ generadas a partir de dos fGns distintos con exponentes de Hurst H_1 y H_2 respectivamente, de forma que la serie resultante es $(x(i) = x_{\text{mod}}(i) \cdot x_s(i))$ Gaussiana por naturaleza y presenta correlaciones no-lineales (ver Sec. 2.3). Además, su función de autocorrelación viene dada por:

$$C_x(\ell) = C_s(\ell) \frac{(\pi - 2)C_{|x|}(\ell) + 2}{\pi} \quad (3.100)$$

donde $C_{|x|}$ es la función de autocorrelación de la serie $x_{\text{mod}}(i)$ generada a través del fGn con exponente de Hurst H_1 y C_s es la función de autocorrelación de $x_s(i)$ generada a través del fGn con exponente de Hurst H_2 . Hay que tener en cuenta que aunque la autocorrelación de la serie compuesta no es exactamente una ley de potencias, decae asintóticamente con el exponente de la serie usada para la obtención del signo, H_2 . Además, teniendo en cuenta que la autocorrelación de la serie compuesta está perfectamente descrita por (3.100), que los ruidos fraccionarios Gaussianos tienen una función de autocorrelación dada por 1.12, y las aproximaciones realizadas tanto para la autocorrelación del módulo como la del signo de un fGn Eqs. 3.36 y 3.96, podemos ver su dependencia funcional con el retardo ℓ :

$$C_x(\ell) \simeq \frac{2}{\pi^2} \frac{(2H_2 - 1)2H_2}{2\ell^{2-2H_2}} \cdot \left[\frac{4H_1^2(2H_1 - 1)^2}{4\ell^{2(2-2H_1)}} + 2 \right], \quad (3.101)$$

y finalmente:

$$C_x(\ell) \simeq \frac{2H_2(2H_2 - 1)}{\pi^2 \ell^{2-2H_2}} \cdot \left[\frac{H_1^2(2H_1 - 1)^2}{\ell^{4-4H_1}} + 2 \right]. \quad (3.102)$$

Para valores $0 < H_1, H_2 < 1$ el segundo sumando se impone y, por tanto, la función de autocorrelación de la serie compuesta tiene una dependencia asintótica $C_x(\ell) \propto 1/\ell^{2-2H_2}$, por lo que teniendo en cuenta la Eq. 1.14 podemos ver que las correlaciones de la serie compuesta están controladas asintóticamente por las de la serie signo (ver Figs. 2.2 y 2.6 a).

En la Fig. 3.7 representamos la autocorrelación del módulo en función de la autocorrelación de la serie original para algunos ejemplos de series no lineales generadas mediante la composición de módulos y signos independientes (ver Sec. 2.2). En todos los casos las series han sido generadas a través de un signo con exponente $H_2 = 0.85$, por lo que todas tendrán las mismas correlaciones lineales. En cambio, la serie que da origen al módulo tiene un exponente de correlación H_1 variable, dando lugar a desviaciones respecto al valor esperado para series lineales (Eq. 3.35, línea discontinua). Esto implica que el desacoplamiento de módulo y signo en series independientes siempre conduce a un comportamiento no lineal, o visto de otra forma, en un ruido lineal Gaussiano el módulo y el signo no son independientes pero están acoplados de tal forma que conducen al comportamiento

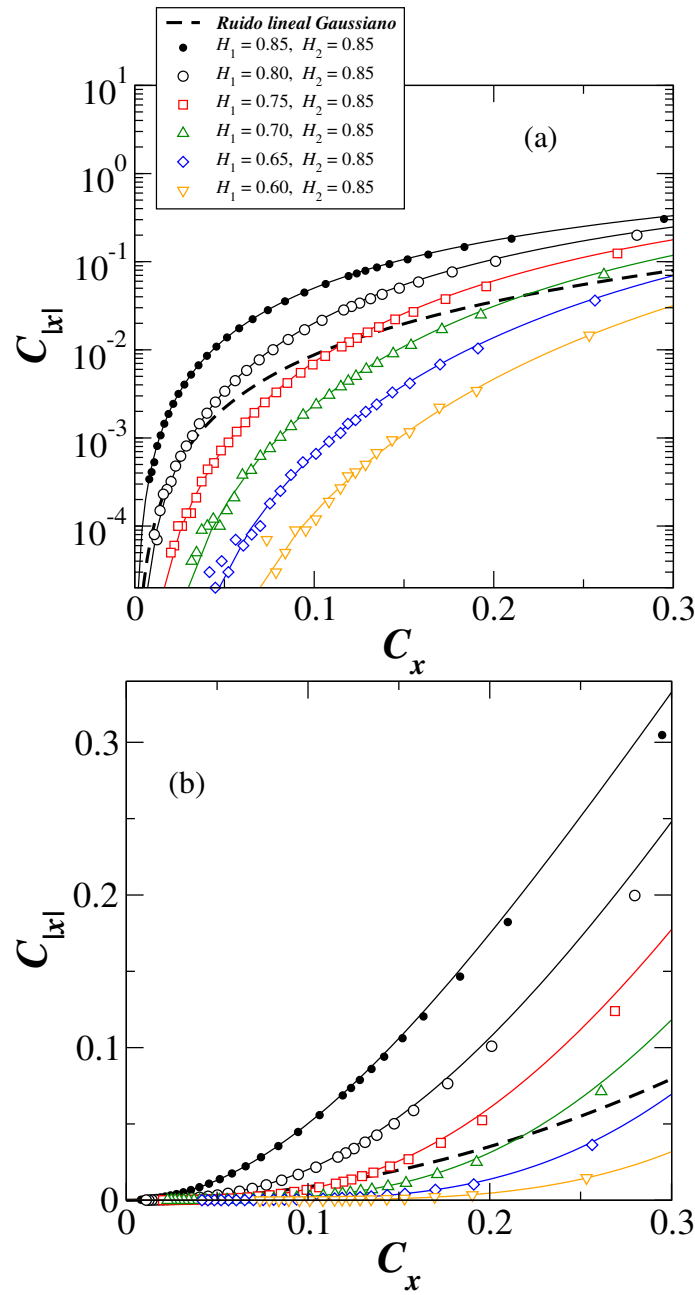


Figura 3.7: (a) Autocorrelación de la serie módulo $C_{|x|}$ en función de la autocorrelación de la serie original, C_x para series no lineales generadas a través de la composición de módulos y signos artificiales (Kalisky et al., 2005) (ver texto). La línea discontinua representa la Eq. 3.35 que ha sido incluida como referencia. Las líneas sólidas corresponden a las curvas teóricas obtenidas usando (3.102), asumiendo que los fGns siguen de forma estricta la Eq. 1.12 para su comportamiento asintótico. Teniendo en cuenta que la Eq. 3.102 es exacta, podemos achacar las desviaciones observadas al hecho de que los fGns creados por el método de Filtrado de Fourier son aproximados, la ecuación para su función de autocorrelación es válida únicamente de forma asintótica y también debido a fluctuaciones estadísticas (sobre todo para pequeños valores de $C_{|x|}$). (b) Mismo gráfico que en (a) con la salvedad de que el eje de ordenadas se encuentra en escala lineal (en (a) se encontraba en escala logarítmica)

mostrado por la Eq. 3.35. También podemos ver (Fig. 3.1) que los ruidos lineales Gaussianos a pesar de presentar correlaciones en su serie módulo, estas van sobre la curva que marca la Eq. 3.35 y por tanto, no presentan desviaciones que indiquen no-linealidad. Debido a los resultados mostrados en la Fig. 3.7 planteamos la posibilidad de considerar la desviación de $C_{|x|}$ frente al valor esperado en el caso de series lineales como una medida de no-linealidad.

3.4.1. No-linealidad y multifractalidad

Multifractalidad y no-linealidad son dos conceptos íntimamente ligados. De hecho, la anchura del espectro multifractal está asociada al grado de no-linealidad de la serie (Badin and Domeisen, 2016; Frisch and Parisi, 1985) y la existencia de propiedades multifractales está relacionada con interacciones complejas no lineales en el sistema bajo estudio. Sin embargo, los resultados obtenidos en la Sec. 2.3 parecen discrepar a priori con los obtenidos en la Fig. 3.7 debido a que mediante el uso de las técnicas de análisis de fluctuación obtenemos valores no nulos para la anchura del espectro multifractal siempre y cuando $H_1 > 3/4$ mientras que en la Fig. 3.7 obtenemos desviaciones frente a la curva del ruido lineal Gaussiano siempre que desacoplamos módulo y signo. Para resolver esta cuestión, en primer lugar definimos un índice de no-linealidad Δ equivalente al promedio del valor absoluto de la desviación cuadrática de la autocorrelación del módulo frente a la esperada en el caso de un ruido lineal Gaussiano desde $\ell = 1$ hasta un valor ℓ_{\max} , i.e.

$$\Delta = \sum_{\ell=1}^{\ell_{\max}} [C_{|x|} - C_{|x|,\text{lineal}}]^2 \quad (3.103)$$

donde $C_{|x|}$ es la autocorrelación del módulo de la serie bajo estudio y $C_{|x|,\text{lineal}}$ es la autocorrelación del módulo en el caso de un ruido lineal Gaussiano. En el caso de series generadas de forma artificial mediante la composición de módulo y signo independientes, el índice de no-linealidad Δ puede ser calculado de manera analítica teniendo en cuenta las Eqs. 2.14 - 3.35 - 3.57 - 1.11. Con objeto de simplificar los cálculos denotaremos:

$$C_{|x|}^{\text{fgn}}(\ell) = \frac{H_1(2H_1 - 1)}{\ell^{2-2H_1}} \quad (3.104)$$

$$C_s^{\text{fgn}}(\ell) = \frac{H_2(2H_2 - 1)}{\ell^{2-2H_2}}. \quad (3.105)$$

Aplicando la Eq. 2.14 y teniendo en cuenta las Eqs. 3.35 y 3.57 obtenemos la autocorrelación de la señal compuesta:

$$C_x(\ell) = \frac{4}{\pi^2} \arcsen(C_s^{\text{fgn}}(\ell)) \cdot \left(C_{|x|}^{\text{fgn}}(\ell) \cdot \arcsen(C_{|x|}^{\text{fgn}}(\ell)) + \sqrt{1 - C_{|x|}^{\text{fgn}}(\ell)^2} \right) \quad (3.106)$$

mientras que la correlación del módulo de la serie compuesta, $C_{|x|}(\ell)$ viene dada por:

$$C_{|x|}(\ell) = \frac{2}{\pi - 2} \left[C_{|x|}^{\text{fgn}}(\ell) \cdot \arcsen(C_{|x|}^{\text{fgn}}(\ell)) - 1 + \sqrt{1 - C_{|x|}^{\text{fgn}}(\ell)^2} \right], \quad (3.107)$$

y su valor esperado en el caso de ruido lineal Gaussiano:

$$C_{|x|, \text{lineal}}(\ell) = \frac{2}{\pi - 2} \left[C_x(\ell) \cdot \arcsen(C_x(\ell)) - 1 + \sqrt{1 - C_x(\ell)^2} \right]. \quad (3.108)$$

Finalmente, reemplazando mediante las Eqs. 3.107 y 3.108 podemos obtener de forma analítica el índice de no-linealidad Δ en función de los parámetros H_1 y H_2 utilizados para la obtención de series compuestas, es decir:

$$\begin{aligned} \Delta(H_1, H_2) &= \frac{2}{(\pi - 2)^2} \sum_{\ell=1}^{\ell_{\max}} (C_{|x|}^{\text{fgn}}(\ell) \cdot \arcsen(C_{|x|}^{\text{fgn}}(\ell)) + \sqrt{1 - C_{|x|}^{\text{fgn}}(\ell)^2} \\ &\quad - C_x(\ell) \arcsen(C_x(\ell)) + \sqrt{1 - C_x(\ell)^2} - 2)^2. \end{aligned} \quad (3.109)$$

En la Fig. 3.8 se representa el índice de no-linealidad Δ (de forma analítica) en función del valor de H_1 para series generadas mediante la composición de módulos y signos independientes. De forma contraria a lo que puede parecer tras una simple inspección de la Fig. 3.7a donde se aprecia desviación sobre el ruido lineal Gaussiano para todos los valores de H_1 , el índice de no-linealidad Δ empieza a crecer de forma considerable para valores $H_1 > 3/4$ independientemente del valor H_2 que marca las correlaciones del signo (podemos apreciar este resultado en la Fig. 3.7b donde vemos las posibles desviaciones sobre la Eq. 3.35 en escala lineal). Este resultado es muy similar al obtenido en la Fig. 2.11 donde se aprecia que la anchura del espectro multifractal (asociada a las propiedades no lineales de la serie) empieza a crecer sustancialmente a medida que la serie módulo toma correlaciones de largo alcance, i.e. $H_1 > 3/4$. Por tanto, a pesar de los problemas que presentan las técnicas de análisis de fluctuación para estimar correlaciones en la serie módulo y serie signo (Carpena et al., 2017), las medidas de no-linealidad obtenidas mediante el uso del MFDFA no se ven afectadas por resultados espurios (de acuerdo con nuestro índice de no-linealidad). Por otro lado, el método que hemos propuesto para la estimación de propiedades no lineales otorga resultados prácticamente idénticos a los obtenidos me-

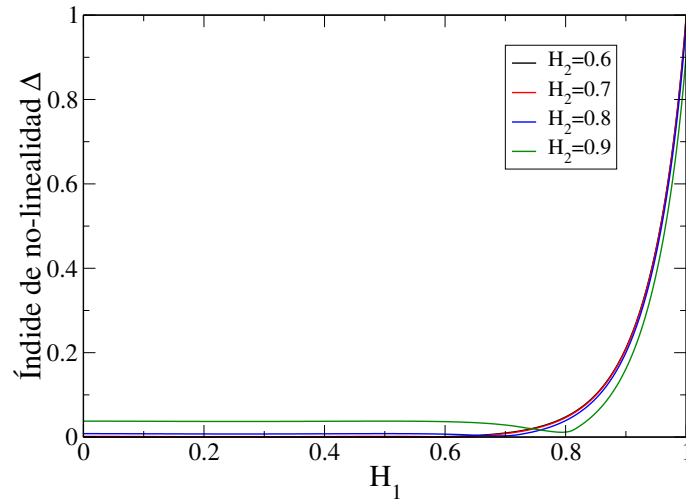


Figura 3.8: Cálculo analítico del índice de no-linealidad Δ a través de la Eq. 3.109 para series artificiales generadas mediante la composición de módulos y signos independientes en función del exponente de entrada para la serie módulo H_1 . Los resultados muestran un claro crecimiento de Δ para valores $H_1 > 3/4$ para todos los valores de H_2 . Este resultado avala los resultados obtenidos en la Fig. 2.11 donde se muestra que la anchura del espectro multifractal (asociada a las propiedades no lineales de la serie) empieza a crecer cuando la serie módulo presenta correlaciones de largo alcance, i.e. $H_1 > 3/4$.

dianete el uso de métodos tradicionales (técnicas de análisis de fluctuación) (Kalisky et al., 2005). Por último, vale la pena mencionar que para valores $H_1 < 3/4$, donde la función de autocorrelación de la serie módulo decae de forma más rápida que $1/\ell$, también existe desviación sobre la curva que rige la Eq. 3.35, por lo que también existirían propiedades no-lineales, aunque el valor del índice de no-linealidad en estos casos es prácticamente despreciable en comparación con los casos donde $H_1 > 3/4$.

3.4.2. Efectos de tamaño finito en medidas de no-linealidad

Los efectos de tamaño finito son siempre un factor a tener en cuenta a la hora de tratar con series temporales. De hecho, en algunos casos, el obviarlos puede conducir a resultados espurios y erróneos. Sin ir más lejos, en el Cap. 2 vimos que en la descomposición, las correlaciones de la serie signo se aproximan a las correlaciones de la serie original a medida que aumentamos el tamaño de la secuencia. También, jugaban un papel esencial en la posición del crossover existente durante la composición cuando $\alpha_{\text{mag}} < 1$ y $\alpha_{\text{sign}} > 1$. En la bibliografía también podemos encontrar varios trabajos que se centran en estudiar los efectos de tamaño finito: en el DFA (Coronado and Carpena, 2005), en el ADN (Peng et al., 1993a), en redes con invariancia frente a escala (Boguná et al., 2004), en transiciones de fase (Borgs and Kotecký, 1992), etc.

En nuestro caso, estudiamos en qué medida los efectos de tamaño finito afectan nuestra media de no-linealidad. En primer lugar vamos a hacerlo mediante la generación de ruidos fraccionarios Gaussianos de diversos tamaños, los cuales son ruidos lineales, por lo que a priori sus puntos deben estar sobre la curva que marca la Eq. 3.35. A pesar de eso, queremos ver los efectos de tamaño que propician las fluctuaciones estadísticas. La Fig. 3.9 muestra los resultados. Podemos apreciar que el índice de no-linealidad decrece a medida que aumenta el tamaño de la secuencia, no se ve afectado por el exponente de correlación del fGn en cuestión, y tiene una dependencia funcional en forma de ley de potencias respecto al tamaño de la secuencia, con exponente de caída aproximadamente igual a la unidad. Por lo tanto, para ruidos fraccionarios Gaussianos, el índice de no-linealidad presente, debido a fluctuaciones estadísticas se comporta como:

$$\Delta \propto \frac{1}{N}, \quad (3.110)$$

donde N es el tamaño de la secuencia. Por otro lado, vamos a estudiar también los posibles efectos de tamaño finito en series no lineales. De nuevo, vamos a utilizar las series obtenidas mediante la composición de módulos y signos independientes (Kalisky et al., 2005). La Fig. 3.10 muestra los resultados obtenidos para series generadas con $H_1 = 0.9$ y $H_2 = 0.5$. El coeficiente de no-linealidad no parece verse afectado de forma clara por el tamaño de la secuencia, ya que oscila en torno a un valor constante, por lo que aún no podemos extraer conclusiones claras sobre el efecto del tamaño de la secuencia en la medida de no-linealidad. Como perspectiva de futuro de este trabajo, sería interesante un estudio más en profundidad sobre este aspecto, mediante la búsqueda de relaciones analíticas que nos permitan entender de una forma más clara cómo el tamaño de la secuencia repercute en la no-linealidad. Por el momento, durante los análisis de no-linealidad de series cardíacas que realizaremos en el próximo capítulo procuraremos, en la medida de lo posible, analizar series del mismo tamaño cuando se comparen los índices de no-linealidad.

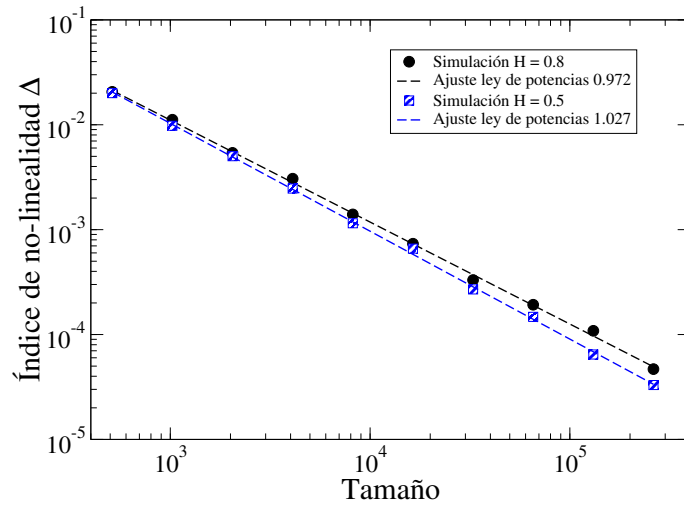


Figura 3.9: Índice de no-linealidad Δ en función del tamaño de la secuencia para fGns con exponente de correlación $H = 0.5$ y $H = 0.8$. Para cada tamaño de secuencia, el experimento se repitió un número de veces de acuerdo con el tamaño de secuencia (más veces para los tamaños menores y menos veces para tamaños mayores), se calculó para cada una de ellas el índice de no-linealidad y posteriormente, el valor medio en cada tamaño. Las líneas discontinuas negras y azules representan el ajuste mediante una ley de potencias para las series con correlación $H = 0.8$ y $H = 0.5$ respectivamente. En ambos casos, se aprecia claramente como el índice de no-linealidad decrece sustancialmente a medida que aumenta el tamaño de la secuencia. Además, Δ parece decrecer de igual forma independientemente del exponente de correlación de la señal en cuestión. Por último, Δ se ajusta de manera óptima a una ley de potencias de forma que $\Delta \propto 1/N$.

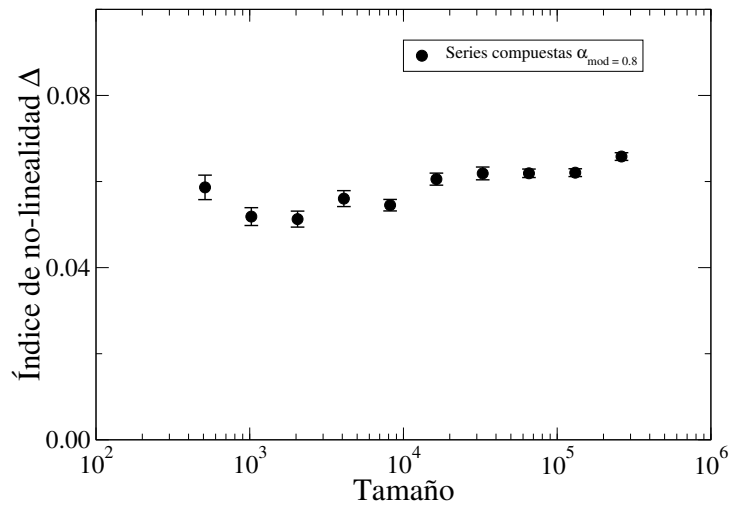


Figura 3.10: Índice de no-linealidad Δ en función del tamaño de la serie, en series generadas mediante la composición de módulos y signos independientes ($H_1 = 0.9$, $H_2 = 0.5$). Para cada tamaño de secuencia, el experimento se repitió un número de veces de acuerdo con el tamaño de secuencia (más veces para los tamaños menores y menos veces para tamaños mayores), indicando los puntos el valor medio y las barras el error estándar. No se puede extraer una clara conclusión sobre los efectos que el tamaño de la serie propicia en el coeficiente de no-linealidad.

Capítulo 4

Medidas de no-linealidad en series cardíacas

En el campo de las series temporales, el concepto de no-linealidad puede ser interpretado de varias maneras (Ashkenazy et al., 2003). Una definición intuitiva consiste en que las series no lineales son aquellas generadas mediante sistemas regidos por dinámicas no lineales, i.e. los valores de la serie temporal dependen de expresiones no lineales tales como cuadrados, logaritmos, funciones trigonométricas, etc. Sin embargo, normalmente no tenemos información alguna sobre esta dependencia funcional, y en muchos casos, el logro es simplemente hallar las ecuaciones que rigen la dinámica del sistema. Otra definición de no-linealidad, aunque no tan intuitiva, es la aportada por Schreiber and Schmitz (Schreiber and Schmitz, 2000), por la que una serie es lineal cuando sus fases de Fourier son aleatorias, i.e. los valores de esta serie están uniformemente distribuidos en el intervalo $[-\pi, \pi]$. Entonces, la presencia de correlaciones no lineales en la serie podría ser evaluada mediante el siguiente test:

- i) Dada una serie temporal, calculamos la transformada de Fourier, aleatorizamos sus fases y calculamos la transformada inversa de Fourier. De esta forma, nos aseguramos que las correlaciones lineales no varían ya que al no modificar la amplitud de la transformada de Fourier, el espectro de potencias no cambia y consecuentemente, tampoco lo hacen las correlaciones lineales (Schreiber and Schmitz, 2000).
- ii) Se realizan tests estadísticos entre la señal original y la señal surrogada obtenida en i) y se comprueba si hay diferencias significativas entre ambas. En el caso de haberlas implicaría que la señal original tiene fases de Fourier no aleatorias y, por tanto, correlaciones no lineales. En caso de no existir tales diferencias, nuestra señal original sería lineal.

En muchos casos, el objetivo es determinar el grado de no-linealidad de la serie bajo estudio, y se usa el valor del propio estadístico como medida de no-linealidad. Las correlaciones en la serie módulo han demostrado ser un buen indicio para estudiar la existencia de correlaciones no lineales (aunque ya vimos que no siempre es así, e.g. series monofractales pueden poseer correlaciones en la serie módulo, y sin embargo, no presentar propiedades no lineales). Una vez obtenida la serie módulo, el procedimiento estándar para evaluar sus correlaciones suele ser el uso del DFA. Como ya hemos visto en el capítulo 3, el uso del DFA induce en ciertos casos a errores en la estimación del exponente de correlación de series módulo y signo (ver de forma detallada en (Carpena et al., 2017)). Además, requiere que la dependencia funcional entre la fluctuación y el tamaño de ventana sea una ley de potencias (existencia de scaling), algo que en muchos casos es una utopía a la hora de estudiar series temporales procedentes de sistemas reales. Por estas razones, en el capítulo 3 propusimos una nueva medida de no-linealidad basada en la desviación que experimentan las correlaciones de la serie módulo bajo estudio, $C_{|x|,exp}$, frente a las que cabría esperar en el caso de un ruido lineal Gaussiano, $C_{|x|}$. A lo largo de este capítulo vamos a aplicar esta nueva medida de no-linealidad al estudio de series fisiológicas (ritmo cardíaco) en diferentes estados y condición.

Desde los primeros trabajos (Peng et al., 1993b), mucho se ha hablado sobre las correlaciones de la serie RR (serie temporal cuyos elementos son los periodos de tiempo que transcurren entre dos latidos consecutivos). De hecho, se ha probado que las correlaciones en esta serie constituyen una poderosa herramienta para detectar alteraciones debidas a patologías o la edad (Goldberger et al., 2002; Ivanov et al., 1996), discriminar entre estados fisiológicos (Ivanov et al., 1999) y evaluar el grado de forma física (Aubert et al., 2003; Dong, 2016). En la mayoría de los casos, los trabajos se limitan al estudio de las correlaciones lineales (espectro de potencias, función de autocorrelación, DFA, etc), pasando por alto la presencia de correlaciones no lineales, a pesar de jugar un papel importante en la dinámica del corazón, ya que su ausencia ha sido asociada a situaciones patológicas y edad avanzada (Ashkenazy, 2001; Ivanov, 1999a).

4.1. Estudio de no-linealidad durante ejercicio y reposo

En cuanto al comportamiento del corazón durante el ejercicio físico es bien sabido que la dinámica cambia de forma ostensible. El cambio más evidente es el incremento abrupto que se produce en el ritmo cardíaco (reducción del valor medio de los intervalos RR) y la reducción de la variabilidad cardíaca (HRV), i.e. reducción de la varianza de

la serie RR (Sarmiento et al., 2013). Junto con estos cambios que pueden apreciarse tras una simple inspección visual de la serie RR, también se ha probado que el ejercicio físico modifica el espectro de potencias de la serie, reduciendo las componentes en baja frecuencia (Anosov et al., 2000; Brodie, 2006; Sarmiento et al., 2013) e introduciendo componentes de muy alta frecuencia, altamente influenciadas por el ritmo respiratorio (Lewis and Short, 2010; Martín-González and García-Manso, 2012), además de reducir la entropía (Platisa et al., 2008). En cuanto a las correlaciones lineales es frecuente utilizar como estimador el exponente de corta escala del DFA, α_1 (exponente de scaling de $F(\ell)$ vs. ℓ en el intervalo comprendido entre 4–11 latidos). En la bibliografía se puede encontrar que las correlaciones lineales de corto alcance están altamente influenciadas por la respiración (Perakakis et al., 2009), aumentan durante el ejercicio físico (Karasik et al., 2002; Platisa et al., 2008) y con la intensidad del ejercicio (A. J. Hautala et al., 2003), aunque el resultado opuesto también puede ser encontrado (Tulppo et al., 2001). Debido a esto, merece la pena comentar que a pesar de que el estimador α_1 ha sido utilizado en cientos de trabajos para cuantificar correlaciones lineales de corto alcance, es un estimador de dudosa eficacia por varias razones:

- i) Tamaño insuficiente de la muestra estadística (realiza una regresión con una muestra de aproximadamente 4 puntos).
- ii) Pasa por alto la correlación a distancia 1-4 (puntos donde la función de autocorrelación está prácticamente exenta de ruido y arroja información muy importante acerca de la dinámica del sistema).
- iii) Está altamente influenciado por el ritmo respiratorio (Perakakis et al., 2009).
- iv) Ausencia de scaling en la serie RR, sobre todo en el transcurso de la actividad física.

En resumen, a pesar de las discrepancias originadas por el uso indebido del estimador α_1 , está claro que el ejercicio físico reduce la complejidad de la serie RR, y que este efecto se debe posiblemente a la rotura del equilibrio presente entre las ramas simpática y parasimpática del sistema nervioso autónomo (SNA) (ver Apéndice C), debido a una retirada del sistema parasimpático y/o activación del sistema simpático (Brodie, 2006; Lewis and Short, 2010).

En esta sección vamos a realizar un estudio sobre la no-linealidad en series RR de reposo y ejercicio en futbolistas aficionados y profesionales. Partimos de la hipótesis de que la reducción en complejidad mencionada previamente debe verse transferida también en una pérdida de no-linealidad. En particular, vamos a centrarnos en el estudio a corta escala ($\ell < 11$ latidos) ya que en este rango los trabajos encontrados en la bibliografía

Suj	Edad	lat/min	SDNN (ms)	SDSD (ms)	NN50 (ms)	pNN50 (%)
1	22	63.51	93.51	71.14	106	16.99
2	22	67.26	82.35	63.20	132	19.56
3	28	88.22	44.97	26.40	20	4.12
4	23	61.51	98.84	72.64	108	18.00
5	23	69.98	111.59	91.50	120	17.14
6	25	68.72	74.83	58.84	102	14.87
7	23	80.72	82.00	33.85	44	5.50
8	23	75.77	95.45	54.67	93	11.83
9	21	66.00	83.08	47.57	75	12.50
10	20	71.71	101.30	50.51	98	13.94
Med	23	71.34	86.79	57.03	89.80	13.45

Tabla 4.1: Medidas de variabilidad caríaca (HRV) durante el reposo en futbolistas aficionados.

afirman que las correlaciones lineales se ven claramente influenciadas por la actividad física. En primer lugar, analizamos registros en reposo y moderada actividad física de 10 futbolistas aficionados, todos ellos hombres (edad 23.0 ± 2.1 años, ver Tabla 4.1) que carecen de patologías cardiovasculares en su historial clínico. Cada registro incluye dos etapas: (i) 10 minutos en reposo, tumbados en posición supina sobre el campo de fútbol (ii) seguidos de 20 minutos de ejercicio moderado (carrera a ritmo de calentamiento, ver Fig. 4.1a). El ritmo cardíaco fue monitorizado latido a latido mediante dispositivos Polar S810i RR (Polar Electro, Oy, Finland) (Weippert et al., 2010).

Debido a que normalmente las series RR no son estacionarias, especialmente durante el transcurso de la actividad física (Figs. 4.1b y 4.1c), es común utilizar la serie de incrementos de RR en lugar del propio RR. Los incrementos de la serie RR se definen mediante:

$$\Delta RR_i = RR_{i+1} - RR_i. \quad (4.1)$$

Los incrementos de la serie RR se caracterizan por tener un comportamiento bastante estacionario (al menos en el sentido débil), siendo sus distribuciones de valores bastante simétricas (ver Figs. 4.1f y 4.1g) y aunque no se ajustan perfectamente a una curva Gaussiana, son distribuciones *Levi-estables* con colas decayendo de forma más lenta que en el caso Gaussiano (Peng et al., 1993b). Debido a esto, es conveniente convertir los datos de forma que tengan una distribución de valores Gaussiana (se debe recordar que en el Cap. 3 se calculó la relación analítica entre la autocorrelación del módulo y autocorrelación de serie original para series temporales con distribución de valores Gaussiana) mediante

la siguiente transformación (Sec. 1.8):

$$x' = \Phi^{-1}[F(x)] \quad (4.2)$$

donde x' denota la serie de datos ΔRR tras la conversión, $F(\cdot)$ es la distribución acumulada de los datos originales ΔRR y Φ es la distribución normal estándar acumulada $\mathcal{N}(0, 1)$. Además, hemos observado que esta transformación prácticamente no altera las correlaciones lineales de la serie bajo estudio (Fig. 1.10). Para cada sujeto calculamos la función de autocorrelación de la serie de incrementos, $C_x(\ell)$, y la de su módulo, $C_{|x|}(\ell)$ para los registros de reposo y ejercicio. En la Fig. 4.2 mostramos los resultados obtenidos para uno de los sujetos (en este ejemplo se aprecian los puntos calculados con un retardo $\ell = 1, 2, \dots, 20$). En general, observamos que C_x alcanza valores similares durante reposo y ejercicio (incluso mayores en ejercicio) (Fig. 4.2b) pero, por otro lado, $C_{|x|}$ es normalmente mayor durante el reposo (Fig. 4.2c). Además, analizando de una forma más detallada la Fig. 4.2a está claro que no solo los valores de $C_{|x|}$ son mayores en promedio en reposo, sino que también los valores de $C_{|x|}$ en ejercicio están más próximos a la línea que representa el valor esperado en caso de ruidos lineales Gaussianos. Por este motivo, y como ya indicamos en el capítulo anterior, la autocorrelación del módulo $C_{|x|}$ no es un buen índice de no-linealidad, pero sí lo es su distancia al valor esperado en caso de ruidos lineales Gaussianos, es decir:

$$\delta C(\ell) = C_{|x|}(\ell) - C_{|x|, \text{lineal}}(C_x(\ell)), \quad (4.3)$$

donde $C_{|x|, \text{lineal}}(C_x(\ell))$ representa la correlación de la serie módulo en un ruido lineal Gaussiano para un valor determinado de C_x . $\delta C(\ell)$ no solo tiene en cuenta el valor de $C_{|x|}$ sino que también considera su diferencia respecto al valor esperado en el caso de series lineales. Por ejemplo $C_{|x|}(\ell = 1)$ alcanza un valor relativamente alto tanto en reposo como ejercicio (Fig. 4.2c), pero, una vez sustraído el valor esperado en caso de serie lineal, $\delta C(\ell = 1)$ es mucho mayor en reposo que en ejercicio (Fig. 4.2d). Para obtener un único número que cuantifique el grado de no-linealidad presente en la serie, y de forma análoga a lo que hicimos en el capítulo anterior, proponemos que nuestro índice de no-linealidad sea fruto de la suma cuadrática de los valores $\delta C(\ell)$:

$$\Delta = \sum_{\ell=1}^{\ell_{\max}} \delta C(\ell)^2 \quad (4.4)$$

Debido a que estamos interesados en las correlaciones de corto alcance y siguiendo el criterio de muchos autores, adoptamos el valor $\ell_{\max} = 10$. Obtenemos que nuestro índice

de no-linealidad Δ es claramente superior en reposo que en ejercicio (Fig. 4.3). Para cada sujeto Δ es mayor en el registro de reposo, algo que también se consigue en promedio, ya que las diferencias de grupo son claras ($p < 10^{-7}$). Sin embargo, debemos tener en cuenta que al tratar series temporales relativamente cortas, la comparación de resultados entre series de distinto tamaño puede conducir a resultados espurios debido a efectos de tamaño finito. En nuestro caso, los registros durante ejercicio son aproximadamente el doble de largos que los correspondientes al reposo, y añadiendo el hecho de que el ritmo cardíaco aumenta drásticamente durante el transcurso de la actividad física, los registros durante ejercicio poseen 4-5 veces más latidos que los correspondientes al reposo. Por este motivo, comprobamos la validez de los resultados obtenidos comparando registros en reposo con registros en ejercicio con el mismo número de latidos: consideramos un sujeto con N_r latidos durante el reposo y un registro en ejercicio con N_e latidos, siendo $N_e > N_r$ y denotamos $n = \lfloor N_e/N_r \rfloor$. Extraemos n ventanas no solapantes del registro de ejercicio, de izquierda a derecha, y otras n ventanas no solapantes empezando por la derecha y acabando por la izquierda¹. A continuación calculamos Δ para las $2n$ ventanas y promediamos los valores para cada sujeto (ver Fig. 4.3). Aunque ahora las diferencias entre reposo y ejercicio son algo menores, todos los valores de Δ durante el reposo están por encima de los correspondientes valores durante ejercicio (incluyendo las barras de error) y la diferencia entre los valores medios de ambos grupos es aún significativa ($p = 3 \cdot 10^{-4}$).

Siguiendo exactamente el mismo procedimiento descrito anteriormente para el caso de futbolistas aficionados, analizamos una base de datos pertenecientes a futbolistas de élite. En este caso, disponemos de una base de datos con registros de 12 futbolistas con edad 22.1 ± 3.4 años (ver Tabla 4.2, consultar Apénd. C para más información sobre las medidas que aparecen en la tabla). La base de datos consta de registros en reposo (de nuevo tumbados en posición supina) y registros durante la realización del ‘yo-yo’ test (Krustrup et al., 2003) en el que la intensidad de la carrera se incrementa progresivamente. En la Fig. 4.5 analizamos a uno de los sujetos. De nuevo, se aprecia que en la curva $C_{|x|}$ vs. C los puntos correspondientes al reposo están mucho más lejos (en comparación con los puntos correspondientes al test de esfuerzo) de la curva que representa a los ruidos lineales Gaussianos, siendo llamativa la presencia de un punto muy anticorrelacionado, pero situado prácticamente sobre la curva lineal. Del mismo modo que en el caso de futbolistas aficionados, la autocorrelación del módulo en reposo está por encima de la correspondiente al test de esfuerzo.

La Fig. 4.5 muestra los resultados obtenidos en cuanto al índice de no-linealidad para todos los sujetos. De nuevo, y de forma análoga al caso anterior (futbolistas aficionados), el

¹Esto se hace con la finalidad de usar todos los datos disponibles

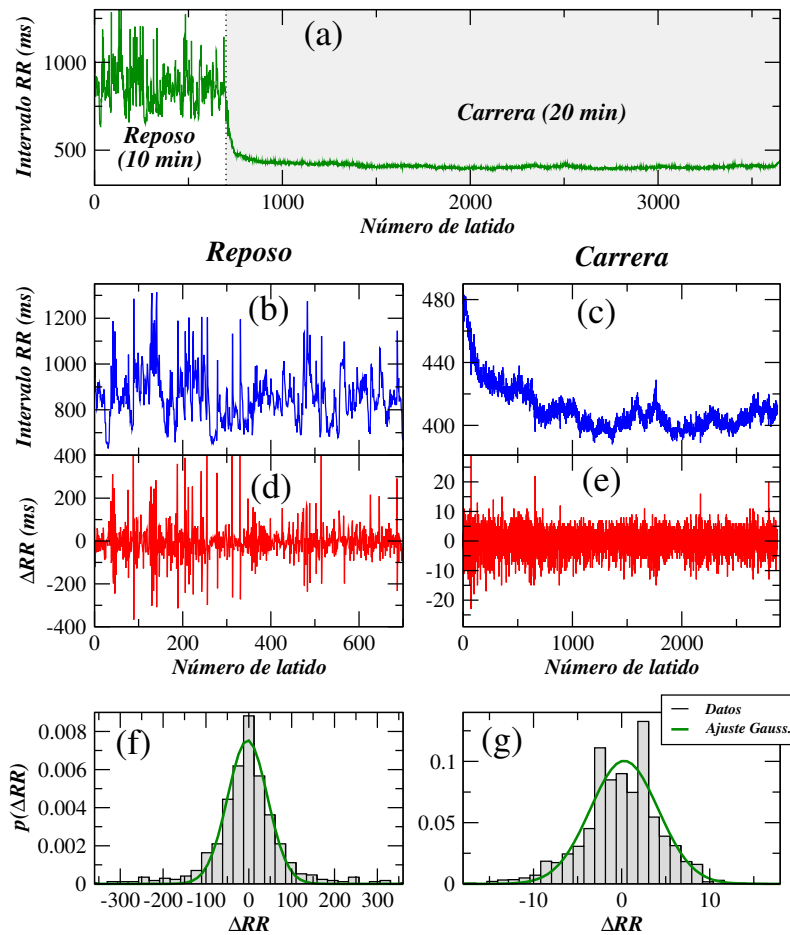


Figura 4.1: Registro RR en reposo y ejercicio moderado procedente de un futbolista aficionado. (a) Registro completo (10 minutos de reposo tumbado en posición supina) y 20 minutos corriendo a un ritmo moderado. (b)(c) Registros separados para reposo y ejercicio respectivamente. (d) (e) Serie de incrementos ΔRR para reposo y ejercicio. (f) (g) Distribución de valores de ΔRR en reposo y ejercicio (se ha incluido el mejor ajuste a una curva Gaussiana).

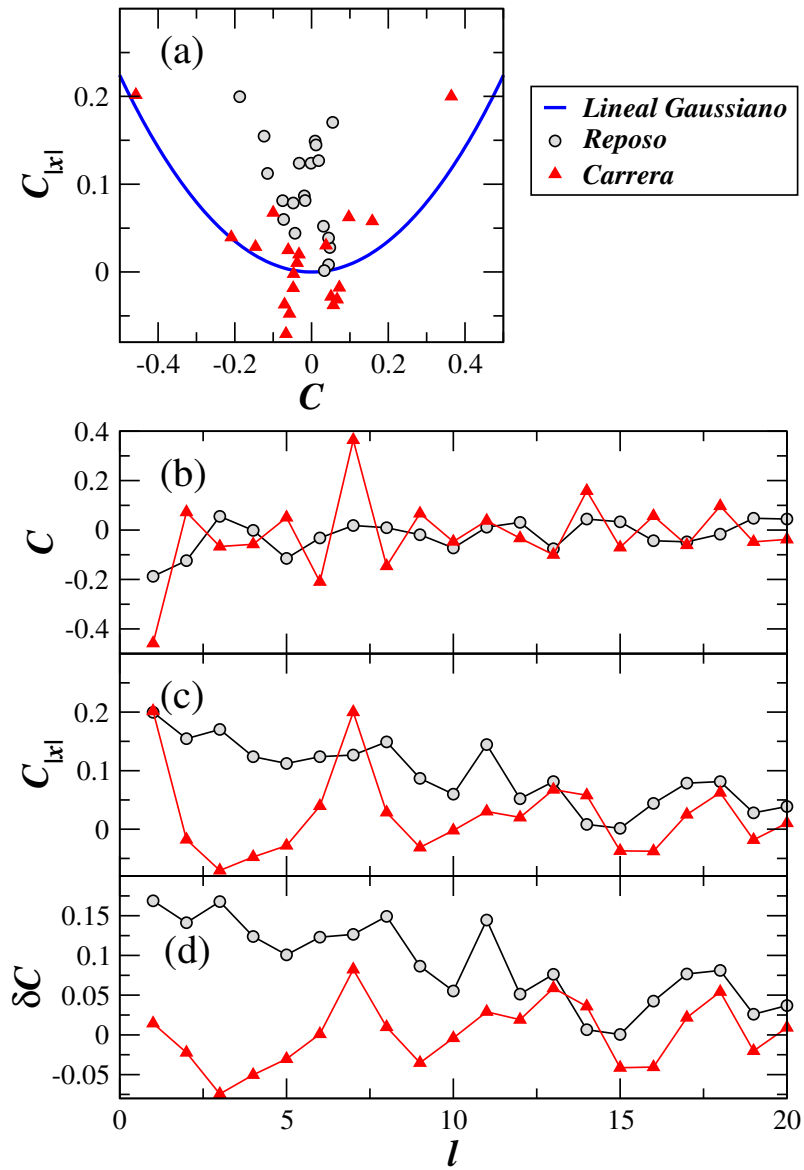


Figura 4.2: Autocorrelación C_x y autocorrelación del módulo $C_{|x|}$ para la serie de incrementos ΔRR durante reposo y ejercicio para el sujeto mostrado en la Fig. 4.1. (a) $C_{|x|}$ vs. C_x durante reposo (círculos) y ejercicio (triángulos). La línea azul corresponde a la curva teórica para el caso de ruidos lineales Gaussianos (Eq. 3.35). (b) Autocorrelación $C_x(l)$ en función de la distancia l durante reposo y ejercicio. (c) Autocorrelación de la serie módulo, $C_{|x|}(l)$ en función de la distancia l durante reposo y ejercicio. (d) Diferencia entre $C_{|x|}$ y el valor esperado para ruidos lineales Gaussianos dado C_x en función de la distancia l durante reposo y ejercicio.

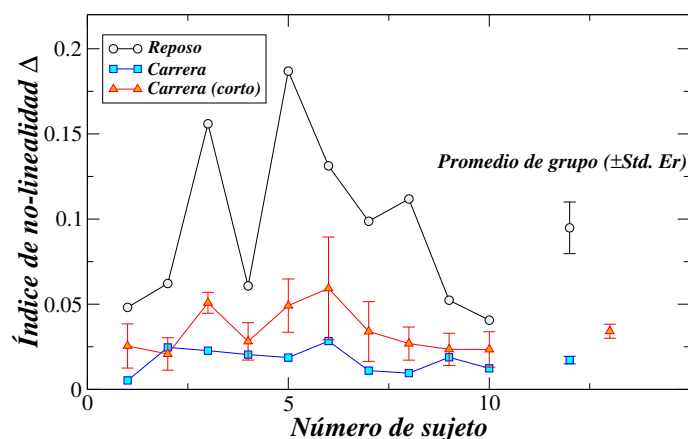


Figura 4.3: Índice de no-linealidad Δ ($\ell_{\max} = 10$) para 10 futbolistas aficionados, hombres, con edad 23.8 ± 2.9 años. Los círculos corresponden a registros de 10 minutos en reposo, tumbados sobre el campo de fútbol. Cuadrados: 20 minutos corriendo a ritmo de calentamiento. Triángulos: promedio de $2n$ subseries (ver texto) obtenidas a partir del registro en ejercicio y con el mismo número de latidos que el correspondiente registro de reposo para cada sujeto (las barras de error representan la desviación estándar sobre el valor obtenido para el coeficiente de no-linealidad en todas las *subseries* de cada sujeto). También se adjuntan los promedios y el error estándar de grupo.

índice de no-linealidad se reduce drásticamente durante el transcurso de la actividad física, existiendo una clarísima diferencia de grupo ($p = 8.6 \cdot 10^{-6}$). También hemos calculado los posibles efectos de tamaño debido a la longitud diferente de los registros en reposo y test, obteniendo en este caso que son prácticamente despreciables (los registros durante el test de esfuerzo tienen una longitud menor que los registros de 20 min de carrera continua usados en el caso de futbolistas aficionados).

Por último, analizamos los promedios obtenidos tanto en futbolistas aficionados y profesionales, en reposo y durante el transcurso de la actividad física. La Fig.4.6 muestra el promedio de cada grupo tanto en reposo como en ejercicio junto con el error estándar. Los futbolistas profesionales muestran en reposo un mayor grado de no-linealidad (en promedio) respecto a los futbolistas aficionados ($p = 4.7 \cdot 10^{-2}$). En cambio, durante el ejercicio físico, los futbolistas profesionales (test de esfuerzo) muestran un menor grado de no-linealidad que los futbolistas aficionados (carrera continua), sin llegar a ser esta diferencia significativa ($p = 5.9 \cdot 10^{-2}$).

A modo de conclusión, el hecho de que el ejercicio físico reduzca de manera drástica el coeficiente de no-linealidad apoya la hipótesis de partida, por la que durante transcurso de la actividad física la señal cardíaca pierde complejidad. Otra conclusión que podemos extraer es el hecho de que, según los resultados obtenidos, existe una relación entre el estado de forma física y el grado de no-linealidad alcanzado en reposo (futbolistas de

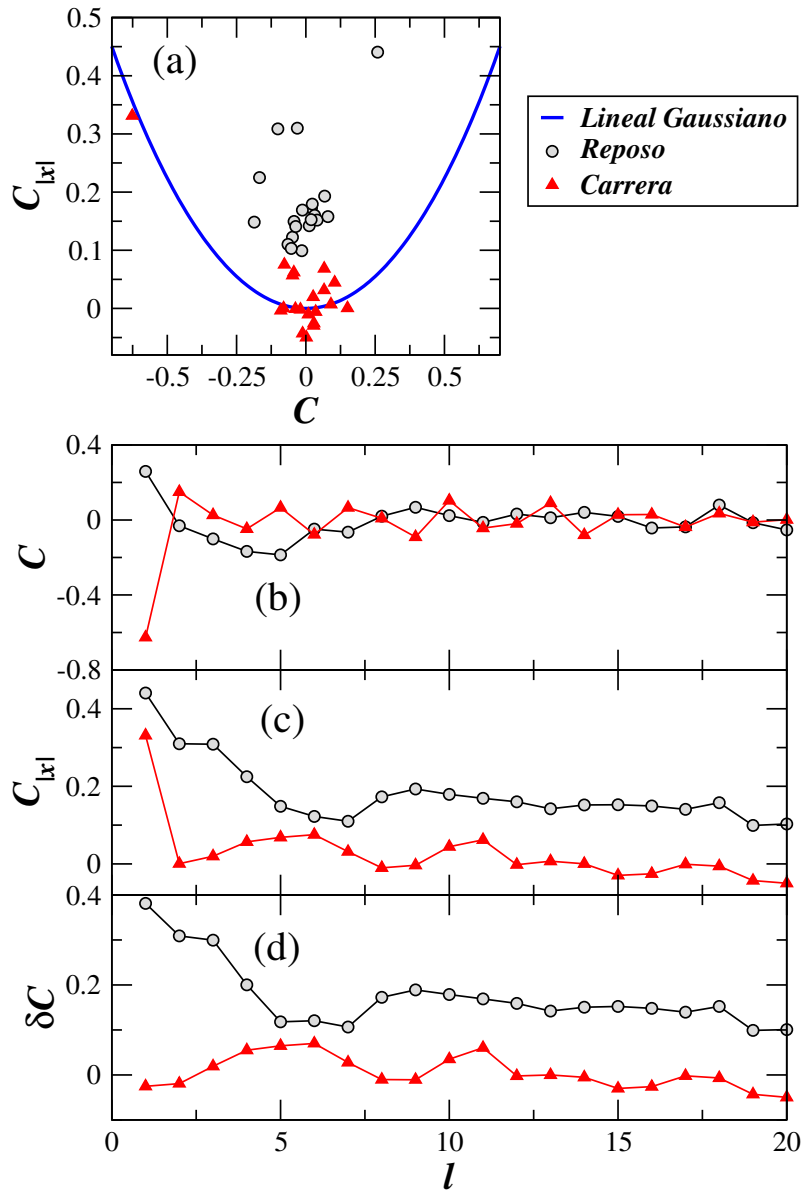


Figura 4.4: Autocorrelación C_x y autocorrelación del módulo $C_{|x|}$ para la serie de incrementos ΔRR durante reposo y test de esfuerzo en un futbolista profesional. (a) $C_{|x|}$ vs. C_x durante reposo (círculos) y ejercicio (triángulos). La línea azul corresponde a la curva teórica para el caso de ruidos lineales Gaussianos (Eq. 3.35). (b) Autocorrelación $C_x(\ell)$ en función de la distancia ℓ durante reposo y test de esfuerzo. (c) Autocorrelación de la serie módulo, $C_{|x|}(\ell)$ en función de la distancia ℓ durante reposo y test de esfuerzo. (d) Diferencia entre $C_{|x|}$ y el valor esperado para ruidos lineales Gaussianos dado C_x en función de la distancia ℓ durante reposo y test de esfuerzo.

Suj	Edad	lat/min	SDNN (ms)	SDSD (ms)	NN50 (ms)	pNN50 (%)
1	28	59.67	123.18	89.75	162	0.269
2	27	64.60	98.21	36.98	44	6.71
3	22	82.00	121.00	49.20	68	8.27
4	18	74.35	118.50	95.60	145	18.19
5	26	63.12	119.73	46.41	47	9.46
6	22	63.02	164.18	109.14	172	23.21
7	20	79.64	97.39	50.29	82	10.95
8	23	71.78	113.68	79.58	114	15.43
9	17	80.19	56.12	27.45	25	3.11
10	20	65.88	94.16	55.91	98	14.80
11	19	62.75	120.63	109.82	173	27.33
12	23	56.26	153.04	101.15	131	24.17
Med	22	68.61	114.99	70.94	105.08	15.79

Tabla 4.2: Datos de futbolistas profesionales durante el reposo.

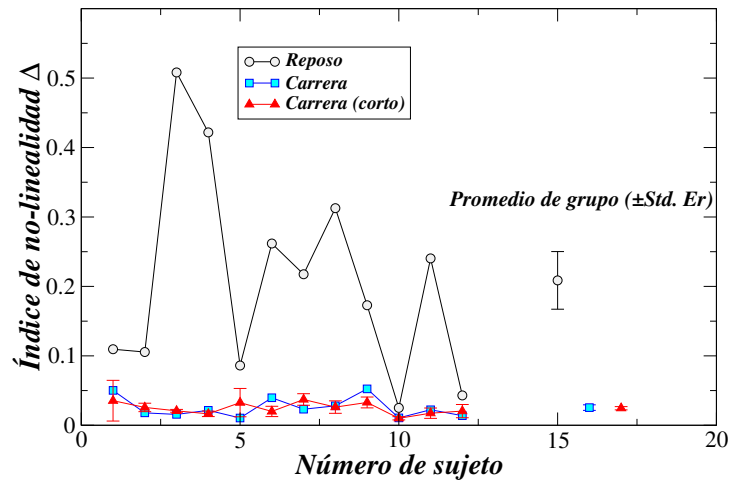
Comparación jugadores profesionales en test

Figura 4.5: Índice de no-linealidad Δ ($\ell_{\max} = 10$) para 12 futbolistas profesionales, hombres, con edad 22.5 ± 3.1 años. Los círculos corresponden a registros de 10 minutos en reposo, tumbados sobre el campo de fútbol. Cuadrados: Realización de un test de esfuerzo con duración aproximada de 5-10 min. Triángulos: promedio de $2n$ subseries (ver texto) obtenidas a partir del registro en ejercicio y con el mismo número de latidos que el correspondiente registro de reposo para cada sujeto (las barras de error indican \pm desviación estándar). También se adjuntan los promedios de grupo y el error estándar.

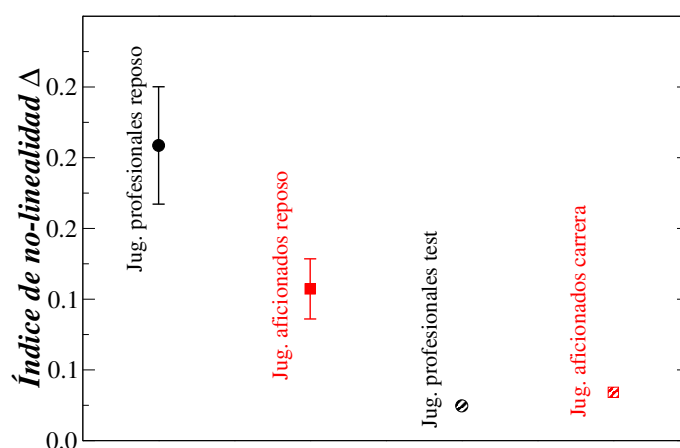


Figura 4.6: Promedio de índice de no-linealidad para futbolistas profesionales en reposo (círculos negros opacos), futbolistas aficionados en reposo (cuadrados rojos opacos), futbolistas profesionales durante test de esfuerzo (círculos negros rayados) y futbolistas aficionados durante carrera continua (cuadrados rojos rayados). Las barras de error representan el error estándar de grupo, y no se aprecian en los casos de ejercicio físico por ser menores que el tamaño de los símbolos empleados. Durante el reposo los futbolistas profesionales muestran un mayor índice de no-linealidad que sus homólogos aficionados ($p = 0.047$). En cambio, durante el transcurso de la actividad física los futbolistas aficionados (carrera continua) presentan un mayor grado de no-linealidad que los profesionales (test de esfuerzo), sin llegar a ser significativa esta diferencia ($p = 0.17$).

élite ostentan un mayor grado de no-linealidad que futbolistas aficionados, $p = 0.047$). También parece existir una correlación entre la reducción de no-linealidad y la intensidad del ejercicio físico, i.e. los futbolistas profesionales presentan menos linealidad durante el test que los futbolistas aficionados (realizan carrera continua a un ritmo suave), a pesar de darse el caso contrario durante el reposo (los profesionales tienen más no-linealidad que los aficionados). En este último caso hemos obtenido un valor $p = 0.17$, por lo que no podemos considerar la diferencia como significativa.

4.2. Estudio de no-linealidad en sujetos con entrenamiento aeróbico y de fuerza

Existen varios trabajos en la bibliografía que intentan mostrar los efectos del entrenamiento en la dinámica del corazón. Se puede encontrar trabajos en los que se estudian las correlaciones lineales a diferentes grados de intensidad de ejercicio (A. J. Hautala et al., 2003), correlaciones lineales en reposo y ejercicio (Karasik et al., 2002), correlaciones lineales en reposo antes y después de someterse a un largo período de entrenamiento (Tulppo et al., 2003), y estudio de la variabilidad cardíaca (HRV) en diferentes situaciones relacio-

nadas con el ejercicio físico (Bernardi et al., 1996; Byrne et al., 1996; Davy et al., 1997; Goldsmith et al., 1992; Levy et al., 1998; Seals and Chase, 1989). En cambio, aunque existen algunos trabajos que estudian cómo varía la complejidad en la señal cardíaca tras someter a los sujetos a periodos de entrenamiento (Heffernan et al., 2007; Kanaley et al., 2009; Tulppo et al., 2001), no existen apenas referencias a la no-linealidad. Los resultados obtenidos hasta el momento son en cierto modo contradictorios, y lo único que se sabe a ciencia cierta es que el ejercicio físico disminuye drásticamente el ritmo cardíaco en reposo y aumenta la variabilidad (basta con un par de semanas de entrenamiento a intensidad baja-moderada en una persona sedentaria para corroborar este efecto).

En esta sección, vamos a estudiar la no-linealidad de la serie RR en reposo de dos grupos de sujetos que realizan habitualmente actividades deportivas totalmente diferentes. Por un lado tenemos a un grupo de culturistas aficionados, formado por 31 hombres, adultos, con edad 28.0 ± 6.1 años (ver Tabla 4.4), y por otro lado tenemos a un grupo de 28 futbolistas (profesionales y aficionados), adultos, con edad 23.0 ± 4.1 años (ver Tabla 4.3). Ambos grupos realizan habitualmente entrenamientos diametralmente opuestos: mientras que los culturistas únicamente realizan entrenamientos de fuerza sin realizar trabajo aeróbico, los futbolistas realizan sesiones de entrenamiento donde predomina la componente aeróbica (sin obviar que la componente anaeróbica siempre está presente en el entrenamiento de fútbol). Para ambos grupos los datos se colectaron mediante dispositivos Polar RS800CX (Weippert et al., 2010), estando los sujetos tumbados en posición supina, y teniendo los registros de ritmo cardíaco una duración aproximada de 10-12 min. La Fig. 4.7 muestra el índice de no-linealidad Δ para cada uno de los sujetos de ambos grupos (el procedimiento para la obtención del índice de no-linealidad es exactamente el mismo que en los casos anteriores). Se aprecia una clara diferencia de grupo, ya que los futbolistas presentan un índice de no-linealidad en promedio mucho más alto que el de los culturistas ($p \simeq 7.6 \cdot 10^{-4}$). Teniendo en cuenta que los sujetos que forman ambos grupos están en un rango similar de edades y sin patologías cardíacas conocidas, podemos asegurar, según los resultados obtenidos, que el entrenamiento aeróbico fomenta la no-linealidad y complejidad en el corazón humano.

4.3. Comparación de resultados mediante el exponente α_1 del DFA

Para mostrar las bondades de nuestra medida de no-linealidad, vamos a comparar nuestros resultados con los obtenidos mediante el uso del exponente α_1 del DFA en algunos de los grupos de datos usados a lo largo de este capítulo. En el caso del DFA, en lugar

Suj	Edad	lat/min	SDNN (ms)	SDSD (ms)	NN50 (ms)	pNN50 (%)
1	28	59.67	123.18	89.75	162	0.269
2	27	64.60	98.21	36.98	44	6.71
3	22	82.00	121.00	49.20	68	8.27
4	18	74.35	118.50	95.60	145	18.19
5	26	63.12	119.73	46.41	47	9.46
6	22	63.02	164.18	109.14	172	23.21
7	20	79.64	97.39	50.29	82	10.95
8	23	71.78	113.68	79.58	114	15.43
9	17	80.19	56.12	27.45	25	3.11
10	20	65.88	94.16	55.91	98	14.80
11	19	62.75	120.63	109.82	173	27.33
12	23	56.26	153.04	101.15	131	24.17
13	22	63.51	93.51	71.14	106	16.99
14	22	67.26	82.35	63.20	132	19.56
15	28	88.22	44.97	26.40	20	4.12
16	23	61.51	98.84	72.64	108	18.00
17	23	69.98	111.59	91.50	120	17.14
18	25	68.72	74.83	58.84	102	14.87
19	23	80.72	82.00	33.85	44	5.50
20	23	75.77	95.45	54.67	93	11.83
21	21	66.00	83.08	47.57	75	12.50
22	20	71.71	101.30	50.51	98	13.94
23	33	64.99	62.23	39.27	54	8.14
24	34	47.26	95.58	116.24	159	33.19
25	28	46.99	91.56	111.13	131	30.61
26	25	77.69	75.83	30.32	34	4.19
27	22	115.56	28.07	8.58	0	0.00
28	23	64.62	119.87	56.20	444	14.73
Med	23	69.82	97.80	63.63	107.29	14.77

Tabla 4.3: Medidas de variabilidad cardíaca (reposo) en futbolistas (profesionales y aficionados).

Suj	Edad	lat/min	SDNN (ms)	SDSD (ms)	NN50 (ms)	pNN50 (%)
1	25	63.21	152.14	137.75	121	17.31
2	40	65.08	74.55	59.49	92	14.07
3	26	63.26	119.62	86.74	156	22.41
4	27	75.39	55.76	38.21	54	7.16
5	22	59.13	68.96	55.75	86	14.43
6	22	56.82	68.25	63.97	82	14.51
7	26	56.01	99.68	86.62	158	28.16
8	20	52.56	126.14	143.25	165	31.19
9	36	60.03	62.58	34.87	30	3.68
10	26	70.02	104.23	100.80	216	28.05
11	20	52.19	182.35	204.67	183	35.12
12	24	54.08	81.13	65.84	110	20.33
13	34	59.87	102.61	65.57	112	18.67
14	30	63.97	60.80	39.69	19	2.97
15	24	59.30	166.94	224.43	131	20.86
16	25	56.04	51.22	52.92	106	18.43
17	24	58.96	68.94	69.15	149	25.25
18	43	60.71	105.45	87.68	191	27.21
19	21	78.79	63.79	39.26	33	4.16
20	35	52.19	182.35	204.67	183	35.12
21	37	64.41	32.40	17.52	2	0.31
22	26	61.60	58.23	29.14	30	4.78
23	27	57.55	130.86	71.90	132	17.84
24	34	49.76	130.89	120.90	124	24.90
25	25	57.88	66.73	70.87	152	25.25
26	26	48.81	182.95	189.24	164	33.40
27	28	48.74	132.44	139.79	131	26.46
28	24	52.76	83.15	72.85	122	20.00
29	35	69.05	53.96	48.17	95	13.73
30	27	59.77	77.05	65.05	129	21.57
31	29	82.52	63.06	29.22	26	2.96
Med	28	60.34	97.07	87.61	112.39	18.72

Tabla 4.4: Medidas de variabilidad cardíaca (reposo) en culturistas (aficionados).

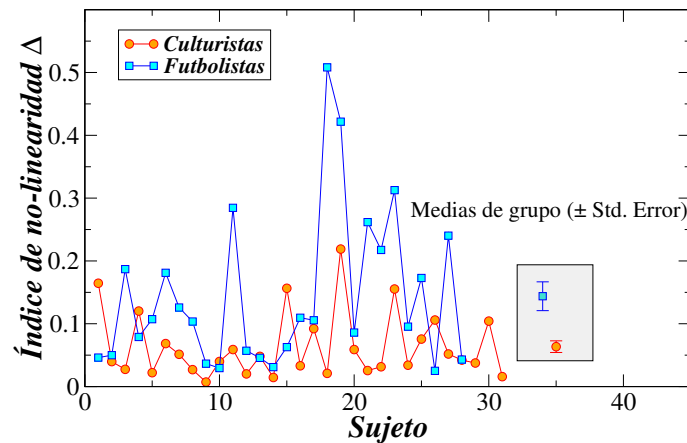


Figura 4.7: Índice de no-linealidad Δ de la serie ΔRR calculado mediante la Eq. 4.4 con $\ell_{\max} = 10$ para futbolistas (profesionales y aficionados) (cuadrados azules) y culturistas (círculos naranjas). Existe efecto de grupo ya que en promedio el índice de no-linealidad es mayor para los futbolistas ($p \simeq 7.6 \cdot 10^{-4}$). En ambos casos, los registros corresponden a series de reposo, tomados con el sujeto tumbado en posición supina y con una duración aproximada de 10-12 min. Las barras de error representan el error estándar.

de analizar la serie ΔRR analizamos la propia serie RR usando previamente un filtro que permite eliminar tendencias de segundo orden, por aquello de la alta no estacionariedad durante el transcurso de la actividad física. El exponente α_1 es una medida que analiza las correlaciones lineales a corta escala de una serie temporal (usualmente entre 4-11 latidos en el caso de la señal cardíaca) y que se utiliza de manera estándar en la bibliografía para discernir entre grupos de datos en distinto estado o condición fisiológica. Como ejemplos relevantes podemos mencionar: sujetos sanos y angina de pecho (Mäkikallio et al., 1998), sujetos sanos y enfermedad coronaria (Laitio et al., 2000), evolución en función de la edad (Pikkujämsä et al., 1999), sujetos sanos y fallo congestivo cardíaco (Peng et al., 1995), reposo y ejercicio (Karasik et al., 2002; Tulppo et al., 2003), ejercicio a diferentes intensidades (A. J. Hautala et al., 2003), etc. En la Fig. 4.8 se muestran los resultados: en primer lugar, en cuanto a los futbolistas aficionados obtenemos un mayor exponente de correlación de corto alcance durante el transcurso de la actividad física, llegando a ser significativa esta diferencia. En el caso de los futbolistas profesionales, aunque la media durante el test de esfuerzo es ligeramente superior a la correspondiente durante el reposo, la diferencia no es ni mucho menos significativa ($p = 0.87$). Por último, en cuanto a la comparación en reposo de futbolistas (tanto profesionales como aficionados) y culturistas aficionados, vemos que los primeros presentan una mayor correlación de corto alcance, siendo significativa la diferencia. Existe cierta contradicción entre los resultados encontrados: si bien en el caso de los futbolistas aficionados el ejercicio parece estimular el exponente α_1 de acuerdo con los resultados mostrados en (Karasik et al., 2002; Platasa

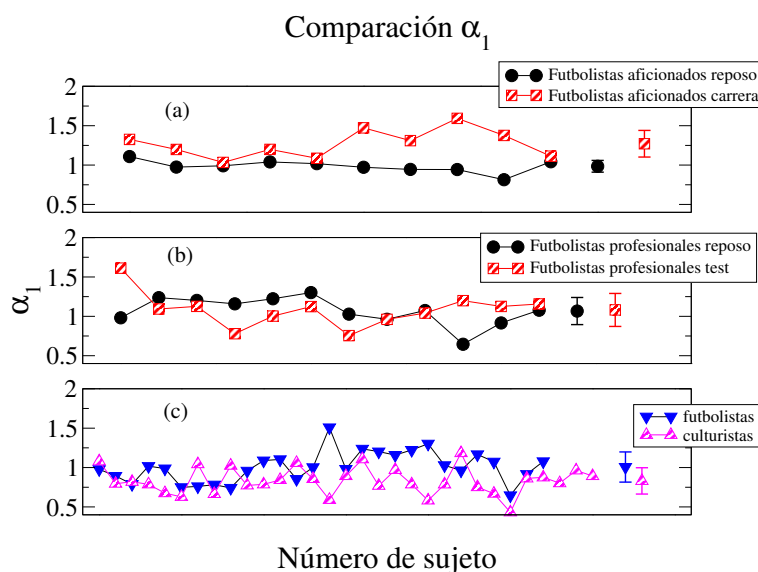


Figura 4.8: Comparación de datos fisiológicos mediante el exponente α_1 del DFA. En este caso el DFA se aplica directamente a las series RR, que previamente han sido sometidas a un filtrado que es capaz de eliminar tendencias de segundo orden. (a) Comparación entre futbolistas aficionados en reposo (círculos negros) y carrera moderada (cuadrados rojos). En la parte derecha del gráfico se adjuntan las medias y desviaciones estándar de grupo, teniendo en este caso que durante la carrera α_1 es mayor que en reposo ($p = 10^{-4}$). (b) Comparación entre futbolistas profesionales en reposo (círculos negros) y durante la realización de un test de esfuerzo (cuadrados rojos). La media durante el test de esfuerzo es ligeramente superior a la obtenida durante el reposo, sin llegar a ser estos resultados significativos ($p = 0.87$). (c) Comparación entre grupo de futbolistas (aficionados y profesionales) (triángulos azules) y culturistas aficionados (triángulos margenta). El valor promedio de α_1 en los futbolistas es mayor que el correspondiente a los culturistas, siendo significativa la diferencia ($p = 5 \cdot 10^{-4}$).

et al., 2008), para los futbolistas profesionales que realizan un test de esfuerzo (mucho más intenso que la carrera moderada) las diferencias con el reposo son insignificantes, en contraposición con lo visto en (A. J. Hautala et al., 2003). Por otro lado, en (Tulppo et al., 2003) hemos encontrado que el entrenamiento aeróbico disminuye levemente el valor del exponente α_1 . Sin embargo, en nuestra comparación entre futbolistas y culturistas observamos el caso contrario: los futbolistas (realizan entrenamiento aeróbico) presentan un mayor valor de α_1 que los culturistas (no realizan entrenamiento aeróbico).

Por tanto, los resultados obtenidos mediante el análisis del exponente α_1 son un tanto contradictorios y no podemos extraer conclusiones claras de los mismos.

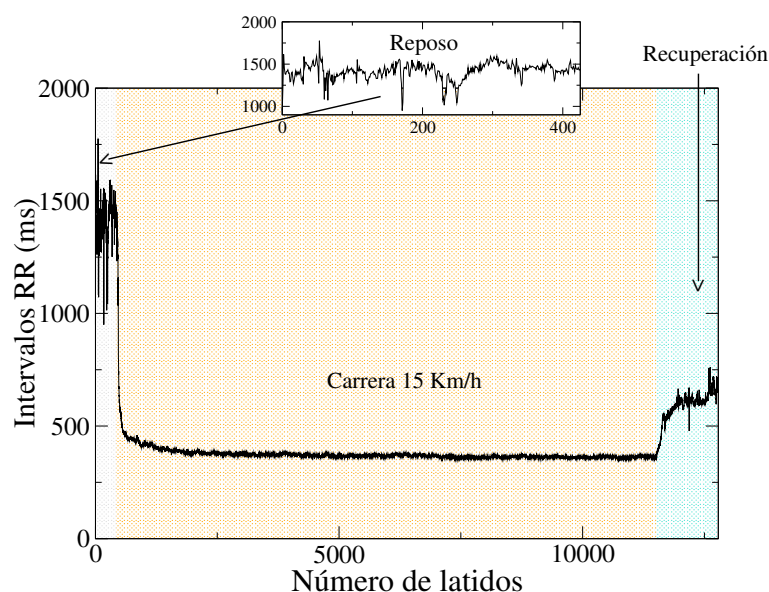


Figura 4.9: Ejemplo de subida y bajada del ritmo cardíaco en un registro de 10 min de reposo (sentado), 50 min de carrera y un período de recuperación en torno a 10 min (caminando a un paso muy lento al principio, y sentado después de los 5 primeros minutos). Se aprecia claramente que justo al iniciar la carrera se produce un aumento brusco del ritmo cardíaco (descenso de los períodos RR). Por el contrario, al iniciar el período de recuperación, el descenso del ritmo cardíaco (aumento de los períodos RR) es mucho más lento.

4.4. Otros resultados preliminares

En esta sección exponemos otros resultados preliminares que hemos obtenido a través de medidas de no-linealidad en series cardíacas. En concreto, hemos obtenido varios registros de un mismo sujeto que constan de carrera continua con una duración aproximada 45-50 min, seguidos de un periodo de recuperación o vuelta a la calma donde el sujeto se encuentra de pie caminando lentamente. Como antesala de los resultados sabemos que los aumentos del ritmo cardíaco (propiciados por la rama simpática del NSA) son mucho más abruptos, que los descensos (propiciados por la rama parasimpática del NSA). A modo de ejemplo gráfico, en la Fig. 4.9 visualizamos un registro cardíaco correspondiente al autor de este trabajo con una duración aproximada de 10 min de reposo (sentado), 50 min de carrera a una velocidad de 15 km/h y un período de recuperación post-carrera aproximado de 10 min. Como podemos apreciar, una vez iniciada la carrera, la rama simpática del Sistema Nervioso propicia un aumento brusco del ritmo cardíaco (descenso de los períodos RR). Sin embargo, una vez iniciada la recuperación, el ritmo cardíaco decrece de una forma mucho más paulatina (aumento de los periodos RR) (ver Apéndice C).

Una vez analizado este ejemplo, parece claro que las propiedades estadísticas de series

cardíacas durante el reposo, ejercicio y recuperación son totalmente diferentes. Además, el período de recuperación ha sido como un indicador del estado de forma física del sujeto, e.g. en el caso de un sujeto que realiza entrenamiento aeróbico de forma habitual y un sujeto sedentario, si ambos se someten a un ejercicio físico de idéntica intensidad y duración, una vez concluido este, el sujeto con mejor estado de forma física logrará reducir su ritmo cardíaco (volver a la calma) mucho antes que el sujeto sedentario. Un completo análisis de las propiedades estadísticas de la fluctuación cardíaca inmediatamente después de la actividad física puede ser encontrado en (Javorka et al., 2002). También ha sido estudiado el período de recuperación como un predictor de mortalidad (Cole et al., 1999), en situaciones patológicas (Imai et al., 1994; Lipinski et al., 2004), su relación con el SNA (Savin et al., 1982), etc. Debido por tanto al interés que puede suscitar el estudio de las propiedades de la serie cardíaca durante el periodo de recuperación tras realizar una actividad aeróbica, nos proponemos comparar el coeficiente de no-linealidad durante el ejercicio y el período de recuperación. Disponemos de 18 registros pertenecientes al Dr. Pedro A. Bernaola Galván tomados durante un periodo de 3 años (edad comprendida entre 44-47 años) y sin que se daten patologías cardíacas en su historial clínico. Cada registro consta de una hora de carrera a velocidad constante seguido de un período de recuperación de 10-15 min (andando a paso muy lento). Los resultados se muestran en la Fig. 4.10. El protocolo aplicado es exactamente el mismo que el expuesto en las secciones anteriores. De nuevo, apreciamos que el índice de no-linealidad es mucho más elevado en promedio durante el ejercicio que el correspondiente durante el período de recuperación post-carrera ($p = 2.8 \cdot 10^{-2}$).

Como ampliación de este estudio, en el futuro buscaremos obtener bases de datos cardíacos de sujetos en diferente estado de forma física, mientras realizan ejercicio aeróbico y con posterior período de recuperación. El ejercicio aeróbico se realizará preferiblemente en una bicicleta estática para poder controlar la intensidad mediante potenciómetro y garantizar que la recuperación se dé exactamente en las mismas condiciones para todos los sujetos. Con este experimento podremos determinar si existe correlación entre el número de latidos recuperados por minuto en cada atleta y su coeficiente de no-linealidad.

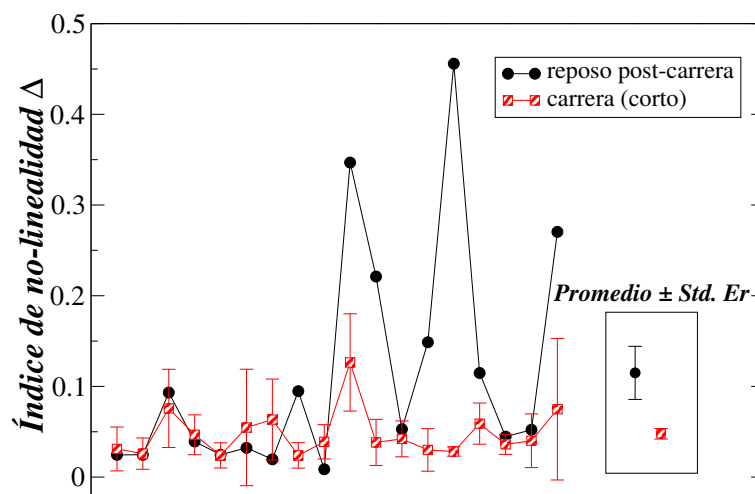


Figura 4.10: Comparación del índice de no-linealidad durante el ejercicio y la posterior recuperación. En el caso de la carrera, para mitigar los posibles efectos de tamaño finito se obtiene el coeficiente de no-linealidad en series del mismo tamaño que el correspondiente a la recuperación, representando los cuadrados el valor medio y las barras de error la desviación estándar sobre todas las series obtenidas a partir del mismo registro. También se adjuntan las medias y errores estándar de grupo, quedando patente que la no-linealidad durante el ejercicio es menor en promedio que la correspondiente al período de recuperación ($p = 2.8 \cdot 10^{-2}$)

Capítulo 5

Conclusions

5.1. Chapter 1

We have done a general review about concepts related to time series, models to reproduce some time series present in nature (fractional Gaussian noise and fractional Brownian noise), methods to generate synthetic time series (Fourier Filtering Method), estimating linear correlations (fluctuation analysis techniques), nonlinear correlations (multifractal spectrum, wavelet transform) and some other useful techniques to deal with time series.

5.2. Chapter 2

In this chapter we have presented a systematic study of the correlation properties (by means of DFA) of the decomposition of artificial long-range power-law linearly correlated time series into their magnitude and sign series as well as the correlation properties, including nonlinear ones, of the composed series obtained as products of independent magnitude and sign series.

Regarding the decomposition problem, we have studied the correlations of the magnitude and sign of a variety of fractional Gaussian noises and fractional Brownian motions generated by means of the Fourier Filtering Method, one of the most widely used to generate artificial linear correlated series. The results are summarized in Fig. 2.3. In addition, we have obtained analytic expressions for the correlation exponent of the sign series α_{sign} (Eq. 2.10), in particular, we show that $\alpha_{\text{sign}} \leq 3/2$ independently of the correlations of the original series. These results, together with those obtained numerically for the magnitude shown in Fig. 2.3 (also in agreement with (Kalisky et al., 2005) for the square of the series), will be of great help in order to model surrogate time series. For example, the sign series obtained from the decomposition are often used to generate correlated binary

series in the study of DNA sequences (Bernaola-Galván et al., 2002, 2012; Peng et al., 1994b) or disordered binary solids (P. Carpena and Stanley, 2002) as well as to generate distributions of first-passage times of correlated series (Carretero-Campos et al., 2012). It is also worth mentioning that, following the results shown in Sec. 2.1 it is clear that long-range anticorrelated binary sequences cannot be obtained using this method, a drawback shared with other methods (Izrailev et al., 2007; Keshet, 2004; Usatenko et al., 2014).

Apart from the utility of the decomposition to generate surrogate series, the comparison of the results obtained here for artificial linear series with those obtained from real data would help to unveil the existence of coupling in the mechanisms responsible for the magnitude and sign or to discard it. This information is instrumental for the study of the underlying processes generating complex non-linear time series such as those obtained from physiological systems.

By means of the composition, we studied the correlations in series obtained as the product of independent series of correlated magnitudes and signs. First, we explore the linear correlations measured by the DFA exponent and find that, only for those composed series with $\alpha_{\text{sign}} < 1$ we obtain a scale free behavior, i.e. a fit to a single power-law of $F(\ell)$ in the whole range. In addition, the correlation exponent of the composition is given by α_{sign} independently of α_{mag} . On the other hand, for $\alpha_{\text{sign}} \geq 1$, we observe clear crossovers for $\alpha_{\text{mag}} < 1$ whose position, ℓ_c , can be obtained analytically (Eq. 2.21). Here the composed signal scales with $\alpha_1 \simeq \alpha_{\text{mag}}$ for $\ell < \ell_c$ whereas for $\ell > \ell_c$ it scales with $\alpha_2 \simeq \alpha_{\text{sign}}$. For $\alpha_{\text{mag}} \geq 1$ we show that crossovers, although theoretically predicted, are difficult to detect in practice and the composed signal approximately scales with a single exponent given by $\alpha_1 = \min\{\alpha_{\text{mag}}, \alpha_{\text{sign}}\}$. As a consequence of this, and taking into account that $\alpha_{\text{sign}} \leq 3/2$, the composition cannot produce signals with correlation exponents above $3/2$. Results are summarized in table 2.1.

Finally, we analyze the nonlinear properties of composed signals by means of MF DFA in the region $\alpha_{\text{sign}} < 1$. As a measure of the nonlinearity in the signal we use the width of the multifractal spectrum ($\Delta\zeta$) and show that it grows almost linearly with α_{mag} , thus indicating that the nonlinear properties of the composed signals are controlled by the correlations in the magnitude. In addition, we also find that $\Delta\zeta$ is independent of α_{sign} . This last result is interesting because it means that we can generate surrogate signals for which we can fix both the linear correlations (α_{sign}) and the strength of the nonlinearities ($\Delta\zeta$).

At last, we would like to remark that some of the results obtained by means of DFA for magnitude and sign time series throughout this chapter are spurious because, just like we show in Chap. 3, DFA is not able to detect correlations in the magnitude series when

$\alpha_{\text{in}} > 3/4$, and in the sign series when $\alpha_{\text{in}} > 1/2$ (Carpena et al., 2017).

5.3. Chapter 3

In this chapter we have obtained the analytic expression between the autocorrelation function of a fractional Gaussian noise, C_x , and its magnitude, $C_{|x|}$ (Eq. 3.35). By visual inspection of $C_{|x|}$ vs. C_x (Fig. 3.1) it is clear that the autocorrelation function of the magnitude time series of a fractional Gaussian noise cannot be anticorrelated, i.e. take negative values. Furthermore, considering the first non-vanishing term of the Taylor expansion of Eq. 3.35 we can see that the autocorrelation function of the magnitude, $C_{|x|}$, shows a quadratic dependence (approximately) with the autocorrelation function of the original series. Moreover, we have also obtained the analytic expression between the autocorrelation function of the quadratic series (Eq. 3.55), C_{x^2} , as a function of the autocorrelation function of the original signal, C_x , also obtaining a quadratic dependence. This results implies that fractional Gaussian noises and their quadratic series have similar asymptotic behavior and justifies that in some papers, scaling properties of magnitude and quadratic time series are associated (Kalisky et al., 2005).

For the sign series (also of great interest when studying important properties of the system under study) exists an analytic relation between its autocorrelation function, C_s , and the original series, C_x , originally found by (Apostolov et al., 2008) for Gaussian series (Eq. 3.57). Throughout this chapter we prove this analytical relation by means of two different ways. Moreover, we show that Eq. 3.57 is not only valid for Gaussian distributions, but for symmetrical distributions. By visual inspection of Eq. 3.57 we see that sign series can be anticorrelated, something that Fluctuation Analysis Techniques are not able to detect (Carpena et al., 2017). Moreover, considering only the first non-vanishing term of the Taylor expansion of Eq. 3.57, we obtain that the autocorrelation function of the sign series, C_s , shows a similar asymptotic behavior to the autocorrelation function of the original series, C_x . This result also corroborate those obtained in Chapter 2 where the scaling exponent of a fractional Gaussian noise takes the same value that the one of the original signal for values $0.5 \leq \alpha_{\text{in}} < 1$ (this is the range where (i) it is possible to calculate the autocorrelation function and (ii) Fluctuation Analysis Techniques do not present spurious results).

In addition, combining Eqs. 3.35 and 3.57 we can obtain an analytic expression which relates the autocorrelation of the magnitude and sign series in linear Gaussian noises (Eq. 3.98), which can be of great utility when studying dynamical properties in systems where the magnitude and sign play an essential role.

Finally, we note that, up to now, all the analytic relations previously mentioned are only valid for linear Gaussian noises. Then, we have obtained nonlinear synthetic series by means of the composition method, and we have checked that in all cases the autocorrelation of the magnitude series shows deviation from the expected value in case of linear Gaussian noises (Eq. 3.35). For this reason, we propose a new measure on nonlinearity based on this deviation. Lastly, we have obtained analytic expressions for the nonlinearity index in synthetic series generated by means of the composition method. The results obtained here are in perfect agreement with those obtained by using the Fluctuation Analysis Techniques in Chapter 2 where the nonlinear properties appear only when the magnitude series is generated with values $H_1 > 3/4$.

5.4. Chapter 4

In this chapter we apply this new measure of nonlinearity to real data (heartbeat series). First, we do a general review of the use of correlations in order to find out dynamical properties in the human heart, e.g. correlations change dramatically with age, physical shape, and different physiological states. However, most of the papers found in the literature are focused on the linear correlations (specially using the short term DFA exponent α_1), being the nonlinear correlations of great importance though, e.g. nonlinear correlations are totally different in healthy and pathological situations. Here, we use the nonlinearity index, Δ , to assess the strength of nonlinearities in:

- **Rest and exercise.** We analyze recordings from amateur soccer player resting and running at a moderate pace and found that nonlinearity is always higher during the rest stage (this result is found for each subject and also in average). We also analyze recordings from professional soccer players during rest and stress test, obtaining also that nonlinearity is dramatically reduced during exercise. This result supports the hypothesis that during exercise the heartbeat time series loses complexity. Furthermore, we also obtain that professional players show higher nonlinearity during rest (in average) than amateur players. This result suggests that the state of fitness (higher for professional players) increases the strength of nonlinearities in the human heart. Lastly, another important conclusion drawn from the analysis is the fact that the intensity of exercise seems also to be involved in the strength of nonlinearities. Professional players show less nonlinearity index during stress test than amateur players running at moderate pace, despite being the opposite during rest, i.e. professional players show higher nonlinearity than amateur players.

- **Cardiovascular and strength training.** Much has been said about the permanent effects of cardiovascular exercise in the human heart. In the literature we can find different results (contradictory in some cases), however, it seems clear that the average cardiac rhythm decreases and heart rate variability (HRV) increases in people who usually do cardiovascular exercise. Here, we calculate the nonlinearity index Δ of recordings from two groups of people who usually follow totally different training methods. On the one hand, we have football players (amateurs and professionals) who are continuously moving during training sessions and thus, cardiovascular component is always present. On the other hand, we have a group of bodybuilders who do not include any cardiovascular activity in their training. Looking at Tables 4.3 and 4.4 we can see similar parameters in both groups, thus, making impossible to draw results from RR statistics. However, our results show that nonlinearity is higher for athletes who include cardiovascular training in the seasons ($p = 7.6 \cdot 10^{-4}$). This result suggests that although nonlinearity is much lower during aerobic episodes, this kind of exercise tends to increase the nonlinearity of the heartbeat time series during rest.
- **Results obtained with DFA exponent α_1 .** Here, we have obtained the α_1 DFA exponent for: (i) rest and exercise episodes from amateur and professional players, finding that α_1 is much higher while running for amateur players. However, despite having been previously published that the intensity of exercise is correlated with α_1 (A. J. Hautala et al., 2003) we find for professional players that α_1 takes practically the same value for rest and stress test. Finally we also compare the results obtained between football players and bodybuilders obtaining that α_1 is slightly higher for soccer players. However, in (Tulppo et al., 2003) we can find the opposite result, i.e. aerobic training decreases the exponent α_1 .
- **Other preliminary results.** Finally, in this chapter we also obtain some new preliminary results. Specifically, we know that recovering periods after exercise are frequently used to assess the state of fitness. Here, we analyze several recordings from Prof. Bernaola-Galván (running and recovering period) and show that, nonlinearity experiments abrupt increase when stopping exercising. This results could be of great interest for the future, as we could draw conclusions about state of fitness just analyzing how the nonlinearity index evolves after exercising.

5.5. General conclusion and future perspective

In this work we have studied correlations properties of magnitude and sign time series by means of DFA and autocorrelation function. We have shown that DFA is not able to accurately estimate correlations in the magnitude and sign time series in some cases. Moreover, we have based on the fact that correlations in the magnitude have widely been used to determine nonlinear properties in time series to propose a new nonlinearity measure which is not affected by the spurious results offered by DFA in some cases. In addition, we have applied this new measure to real data (heartbeat series) during rest and exercise and different training methods.

Finally, as a future perspective, we hope that this work raises awareness of the inappropriate use of DFA found in many papers in the literature and encourages researchers to look for new measures to quantify linear and nonlinear properties in fractal time series. The new nonlinearity index we have proposed here can be applied to analyze nonlinear properties in time series in very different fields and, consequently, better understand the dynamics of complex systems. However, it is still necessary to research how this measure is affected by finite size effects, the ℓ_{\max} parameter and the initial value distribution of the series under study.

Apéndice A

Fluctuaciones del signo en series no estacionarias ($\alpha \geq 1$)

Consideramos una serie con correlaciones de largo alcance con $\alpha \geq 1$ (fBM) y denotamos por x al tamaño de un segmento en cuyo interior no hay ningún cambio de signo (i.e. segmentos con signo constante, o simplemente “segmentos” de ahora en adelante). Se ha probado que la distribución de x sigue una ley de potencias con exponente $\alpha - 3$, la cual una vez normalizada puede ser escrita de la siguiente forma:

$$p(x) = \frac{(2 - \alpha)N^{2-\alpha}}{N^{2-\alpha} - 1} x^{\alpha-3} \quad (\text{A.1})$$

El valor medio $\langle x \rangle$ de los segmentos de signo constante viene dado por:

$$\langle x \rangle = \int_1^N p(x) dx = \left(\frac{2 - \alpha}{\alpha - 1} \right) \frac{N - N^{2-\alpha}}{N^{2-\alpha} - 1} \quad (\text{A.2})$$

y el número medio de tales segmentos en una serie de tamaño N :

$$n = \frac{N}{\langle x \rangle} = \left(\frac{\alpha - 1}{2 - \alpha} \right) \frac{N^{2-\alpha} - 1}{1 - N^{1-\alpha}} \quad (\text{A.3})$$

Al evaluar las fluctuaciones para un tamaño de ventana ℓ , únicamente la porción de señal cubierta por segmentos con $x < \ell$ tendrá contribución no nula: para el resto de la señal, la totalidad de la ventana de tamaño $x < \ell$ estará localizada dentro de un segmento de signo constante y, por tanto, sin fluctuación, por lo que su contribución a $F_{\text{sign}}(\ell)$ será nula¹.

Para evaluar la porción de señal cubierta por segmentos con $x < \ell$, en primer lugar

¹Para ser precisos, por cada segmento de signo constante $x > \ell$ tendremos al menos dos ventanas que pertenecen parcialmente al segmento. Este efecto será prácticamente insignificante para $\alpha > 1$.

evaluamos la probabilidad de que dado un segmento, su tamaño sea más pequeño que ℓ :

$$P(x < \ell) = \int_1^\ell p(x) dx = 1 - \frac{N^{2-\alpha} - \ell^{2-\alpha}}{\ell^{2-\alpha} [N^{2-\alpha} - 1]}, \quad (\text{A.4})$$

el tamaño medio de estos segmentos:

$$\langle x_{<\ell} \rangle = \frac{\int_1^\ell x p(x) dx}{P(x < \ell)} = \left(\frac{2-\alpha}{\alpha-1} \right) \frac{\ell^{\alpha-1} - 1}{1 - \ell^{\alpha-2}} \quad (\text{A.5})$$

y la fracción de serie cubierta por segmentos con $x < \ell$:

$$f(x < \ell) = \frac{n P(x < \ell) \langle x_{<\ell} \rangle}{N} = \frac{\ell^{\alpha-1} - 1}{N^{\alpha-1} - 1} \simeq \frac{\ell^{\alpha-1}}{N^{\alpha-1}} \quad (\text{A.6})$$

Si denotamos por i al número de *unos* en una ventana de tamaño ℓ , se puede obtener de manera directa que la varianza de la ventana viene dada por:

$$\text{var}(i, \ell) = \frac{4i}{\ell} - \frac{4i^2}{\ell^2} \quad (\text{A.7})$$

Teniendo en cuenta que para tamaños de serie N suficientemente grandes encontraremos todos los posibles valores de $i \in \{1, 2, \dots, \ell - 1\}$, podemos asumir que la varianza media en ventanas de tamaño ℓ localizadas dentro de segmentos con $x < \ell$ será:

$$\begin{aligned} \text{var}(\ell) &= f(x < \ell) \langle \text{var}(i, \ell) \rangle_i = \\ &= \frac{\ell^{\alpha-1}}{N^{\alpha-1}} \frac{1}{\ell - 1} \left(\frac{2}{3} \ell - \frac{2}{3\ell} \right) \propto \left(\frac{\ell}{N} \right)^{\alpha-1} \end{aligned} \quad (\text{A.8})$$

y su desviación estándar:

$$\sigma(\ell) = \sqrt{\text{var}(\ell)} \propto \left(\frac{\ell}{N} \right)^{\frac{\alpha-1}{2}} \quad (\text{A.9})$$

$F_{\text{sign}}(\ell)$ mide el valor cuadrático medio de las fluctuaciones de la señal integrada respecto a ℓ y entonces:

$$F_{\text{sign}}(\ell) \propto \sigma(\ell) \cdot \ell \propto \frac{\ell^{\frac{1}{2}(\alpha+1)}}{N^{\frac{1}{2}(\alpha-1)}} = \frac{\ell^{\alpha_{\text{sign}}}}{N^{\alpha_{\text{sign}}-1}} \quad (\text{A.10})$$

donde $\alpha_{\text{sign}} = \frac{1}{2}(\alpha + 1)$ es el exponente DFA de la serie signo para $1 \leq \alpha < 2$.

Para valores más altos de α , las ecuaciones comprendidas entre (A.1) y (A.3) no son válidas (Carretero-Campos et al., 2012). Para estos valores, el número de segmentos n , pasa a ser constante e independiente del tamaño de la secuencia N . Para un tamaño de ventana dado ℓ , únicamente n de las N/ℓ ventanas contribuirán con varianza no nula y,

por tanto, podemos escribir:

$$\text{var}(\ell) = \frac{n\ell}{N} \langle \text{var}(i, \ell) \rangle_i = \frac{n\ell}{N} \frac{1}{\ell-1} \left(\frac{2}{3}\ell - \frac{2}{3\ell} \right) \propto \frac{\ell}{N} \quad (\text{A.11})$$

$$F_{\text{sign}}(\ell) \propto \frac{\ell^{\frac{3}{2}}}{N^{\frac{1}{2}}} = \frac{\ell^{\alpha_{\text{sign}}}}{N^{\alpha_{\text{sign}}-1}} \quad (\text{A.12})$$

donde $\alpha_{\text{sign}} = \frac{3}{2}$ es el exponente DFA De la serie signo para $\alpha \geq 2$.

Se debe tener en cuenta que ambos resultados (Eqs. A.10 y A.12), concuerdan con el hecho de que las fluctuaciones en una serie no estacionaria deben depender del tamaño de la secuencia N .

Apéndice B

Función de autocorrelación de series temporales con módulo y signo desacoplados.

La función de autocorrelación de una serie temporal $\{x_i\}$ a distancia ℓ , cuyos valores siguen una distribución normal con media cero y desviación estándar unidad viene dada por:

$$C(\ell) = \frac{\langle x_i x_{i+\ell} \rangle - \langle x_i \rangle \langle x_{i+\ell} \rangle}{\sigma^2} = \langle x_i x_{i+\ell} \rangle, \quad (\text{B.1})$$

donde $\langle \cdot \rangle$ denota la media sobre la serie temporal. Obviamente podemos escribir:

$$C(\ell) = \langle \text{sgn}(x_i) |x_i| \text{sgn}(x_{i+\ell}) |x_{i+\ell}| \rangle \quad (\text{B.2})$$

$$C(\ell) = \langle \text{sgn}(x_i) \text{sgn}(x_{i+\ell}) |x_i x_{i+\ell}| \rangle, \quad (\text{B.3})$$

donde $\text{sgn}(\cdot)$ denota la fluctuación del signo. Si consideramos que *módulo y signo no están acoplados* (i.e. son variables aleatorias independientes) podemos asumir que:

$$C(\ell) = \langle \text{sgn}(x_i) \text{sgn}(x_{i+\ell}) \rangle \langle |x_i x_{i+\ell}| \rangle \quad (\text{B.4})$$

$$C(\ell) = C_{\text{sign}}(\ell) \langle |x_i x_{i+\ell}| \rangle, \quad (\text{B.5})$$

donde $C_{\text{sign}}(\ell)$ es la función de autocorrelación a distancia ℓ de la serie signo. Por otro lado, en cuanto a la función de autocorrelación del módulo, podemos escribir:

$$C_{\text{mag}}(\ell) = \frac{\langle |x_i x_{i+\ell}| \rangle - \langle |x_i| \rangle \langle |x_{i+\ell}| \rangle}{\langle |x_i|^2 \rangle - \langle |x_i| \rangle^2}, \quad (\text{B.6})$$

y, teniendo en cuenta que $\{x_i\}$ sigue una normal con media cero y desviación estándar unidad, se obtiene que:

$$\langle |x_i| \rangle = \sqrt{\frac{2}{\pi}} \quad \text{y} \quad \langle |x_i|^2 \rangle = 1 \quad (\text{B.7})$$

Sustituyendo en (B.6) obtenemos:

$$\langle |x_i x_{i+\ell}| \rangle = \frac{(\pi - 2)C_{\text{mag}}(\ell) + 2}{\pi} \quad (\text{B.8})$$

y finalmente insertando $\langle |x_i x_{i+\ell}| \rangle$ en (B.4):

$$C(\ell) = C_{\text{sign}}(\ell) \frac{(\pi - 2)C_{\text{mag}}(\ell) + 2}{\pi} \quad (\text{B.9})$$

Apéndice C

Conceptos de Fisiología

Este apéndice ofrece un breve resumen de algunos conceptos que ayudan a entender los mecanismos básicos del funcionamiento del corazón humano. En concreto, vamos a centrarnos en el Sistema Nervioso Autónomo que controla el sistema cardiorespiratorio (ver (Ernst, 2014; Schumann, 2010)).

C.1. Sistema Nervioso Autónomo

El sistema nervioso autónomo (ANS) es una parte periférica del sistema nervioso que controla y regula la homeostasis en el cuerpo humano, i.e. mantiene bajo control ciertas variables fisiológicas como pueden ser la presión sanguínea y el ritmo cardíaco. Cada circuito de control forma una espira cerrada donde el controlador (cerebro o sistema nervioso) obtiene continuamente el valor actual de la variable bajo control, e.g. baro-receptores en arterias miden la presión sanguínea y lo ajustan dependiendo de la situación. Algunas de las propiedades más relevantes controladas por el Sistema Nervioso Autónomo son el contenido de gases, iones y nutrientes en tejidos, salivación, respiración a través de la piel y funciones digestivas y cardiorespiratorias. En este resumen, y debido a los resultados presentados en el Cap. 4 vamos a centrarnos en la parte que atañe a las funciones cardiorespiratorias. El ANS está a su vez dividido en tres partes: (i) sistema simpático (SNS), (ii) sistema parasimpático (PNS) y (iii) el sistema nervioso entérico (ENS). En la regulación de las funciones cardiorespiratorias únicamente intervienen los dos primeros. Ambos componentes actúan de forma complementaria: la activación de la rama parasimpática resulta en la liberación de un neurotransmisor por el nervio neumogástrico que es detectada por el receptor de acetilcolina del corazón que a su vez aumenta la conductividad K^+ en las membranas del corazón y produce la deceleración del ritmo cardíaco. Por el contrario, la aceleración del ritmo cardíaco es inducida por la activación de la ra-

ma simpática debido a la liberación de epinefrina y norepinefrina que llega a receptores adrenérgicos. Diversos estudios han asociado la activación parasimpática con componentes de alta frecuencia ($0.15 - 0.40$ Hz) en la señal de ritmo cardíaco, mientras que la activación de la rama simpática estimula las componentes en baja frecuencia $0.04 - 0.15$ Hz, aunque esto último ofreció mucha controversia ya que algunos estudios apuntan a que la banda de baja frecuencia también es influenciada por la rama parasimpática (Camm et al., 1996; Goldberger, 1999). También se debe tener en cuenta que la activación parasimpática es bastante más rápida debido a una alta concentración de acetilcolinesterasa en el nódulo sinoauricular que induce la contracción del corazón. Sin embargo, la rama simpática predomina en situaciones donde se requiere una rápida respuesta, e.g. levantarse de forma repentina, situaciones en las que hay miedo, etc. Por el contrario, las respuestas a la activación parasimpática parecen mucho más lentas (a modo de ejemplo se puede pensar en el tiempo necesario para volver a la calma después de realizar ejercicio físico). Una posible explicación para esta aparente contradicción radica en la omnipresencia de la componente simpática. La activación simpática puede ser inducida mediante una reducción parasimpática (ocurre con mucha rapidez). Aun así, los mecanismos de regulación del sistema nervioso autónomo no se entienden completamente hoy en día y se siguen investigando.

C.2. Latido Cardíaco

A través de electrocardiogramas (ECG) es posible obtener información sobre el corazón, su dinámica y su estado de salud. El electrocardiograma es una técnica no invasiva que mide las diferencias de potencial inducidas por la excitación cardíaca. Fue introducido por primera vez por Willem Einthoven, quien recibió el premio Nobel en 1924. La Fig. C.1 muestra la configuración de un equipo de medida de electrocardiograma BIOPAC con 3 electrodos, mientras que la Fig. C.2 muestra un típico esquema del ciclo cardíaco en un sujeto sano.

Mecanismos de depolarización y repolarización forman la típica forma de una curva ECG (la nomenclatura PQRST fue introducida por Einthoven). Durante el reposo, las células cardíacas presentan un equilibrio mediante cargas positivas en el exterior de la célula y cargas negativas en el interior, conduciendo esto a una polarización. Una contracción del corazón es entonces simulada por el nódulo sinoauricular, el cual tiene un tamaño de solo unos cuantos milímetros cuadrados. Aunque el nódulo sinoauricular está conectado al sistema nervioso autónomo, y la repetición de sus pulsos puede ser modulada por activación parasimpática y simpática, su función como marcapasos es generalmente

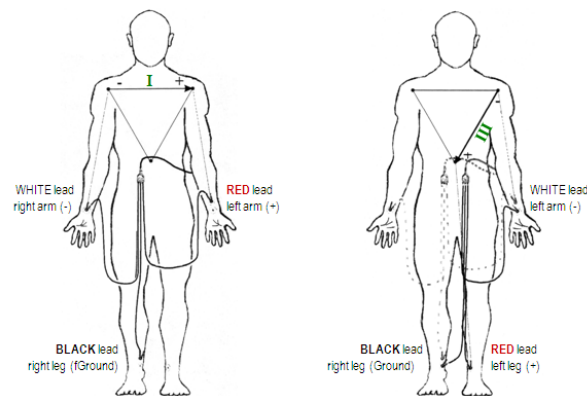


Figura C.1: Configuración de electrodos en un equipo de medida BIO-PAC. Imagen encontrada originalmente en http://tktamop.elte.hu/online-tananyagok/physiology_practical/ch05s02.html

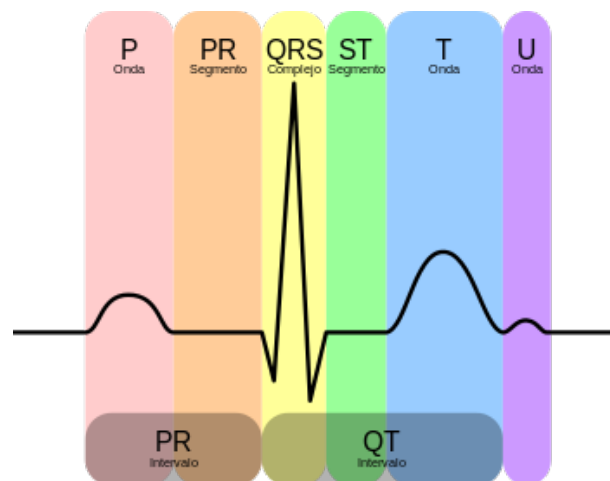


Figura C.2: Ejemplo de ciclo cardíaco. Imagen encontrada originalmente en <https://es.wikipedia.org/wiki/Electrocardiograma>

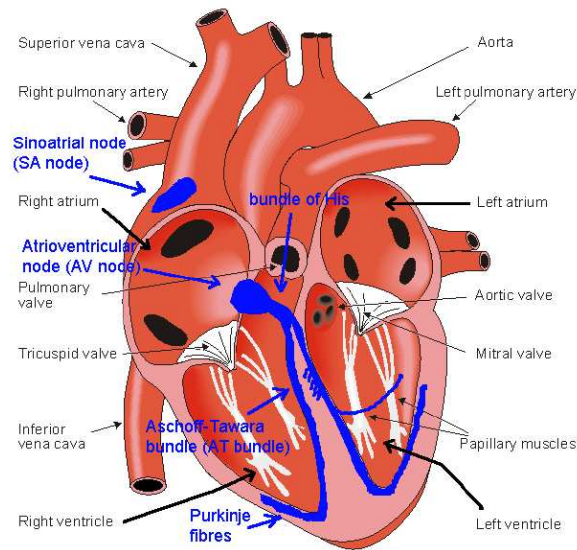


Figura C.3: Esquema del corazón humano. (Imagen publicada originalmente en (Schumann, 2010)).

independiente debido a células auto-excitables y descargas naturales que ocurren a escalas aproximadas de 100 min^{-1} (Despopoulos and Silbernagl, 2003). El nódulo sinoauricular induce la depolarización en la aurícula derecha (aproximadamente 50 ms después de la generación del impulso) la cual evoluciona hasta la aurícula izquierda (85 ms después) y después abajo hasta el nódulo atrioventricular. Esta depolarización de la aurícula se manifiesta en la onda P (ver Fig. C.2). Cuando el pulso llega al nódulo auroventricular es transmitido al haz de His (segmento PQ, llegada - 50 ms, transmisión - 125 ms) desde donde es conducido hasta las fibras de Purkinje. La depolarización de los ventrículos se manifiesta en el complejo QRS, donde el pico R está asociado al ápex del corazón. La repolarización del tejido cardíaco finalmente genera la onda T y el ciclo cardíaco empieza de nuevo. Como acabamos de ver, el ciclo cardíaco es algo complejo. Sin embargo, el pico R es su característica más prominente. Por consiguiente el período RR no es más que el periodo entre dos latidos consecutivos, y la serie temporal RR (estudiada en el Cap. 4) es la serie cuyos elementos son los periodos entre latidos consecutivos.

C.2.1. Observables cardíacos

Con motivo de entender la dinámica del corazón, evaluar el riesgo, e identificar patologías cardíacas que puedan inducir a situaciones peligrosas como el infarto o la muerte súbita, se han propuesto una serie de parámetros de medida. Los más populares son extraídos a partir de la serie RR y se pueden dividir en medidas temporales o medidas

espectrales (Camm et al., 1996; Ernst, 2014). Entre las primeras podemos encontrar la desviación estándar de la serie RR (SDNN), la desviación estándar de los promedios de la serie RR en periodos de 5 min (SDANN), la raíz cuadrada de la media de la suma de los cuadrados de las diferencias entre RRs consecutivos (RMSSD), la desviación estándar de los incrementos de RR (SD SD), el número de pares de RRs consecutivos que difieren en más de 50 ms (NN50), y NN50 dividido entre el número total de intervalos RR (pNN50). En cuanto a los espectrales cabe destacar: (i) la potencia en alta frecuencia (HF, $f \in (0.15, 0.40]$ Hz), (ii) potencia en baja frecuencia (LF, $f \in (0.04, 0.15]$ Hz), (iii) potencia en muy baja frecuencia (VLF, $f \in (0.033, 0.04]$ Hz), (iv) potencia en ultra baja frecuencia ($f \leq 0.0033$ Hz) aunque este último parámetro solo puede ser evaluado en registros de gran duración (~ 24 h).

Apéndice D

Publicaciones y contribuciones

Los resultados de la investigación han dado lugar a varias publicaciones y comunicaciones que se listan a continuación.

D.1. Publicaciones

- M. Gómez-Extremera, P. Carpena, P. Ch. Ivanov and P. A. Bernaola-Galván. Magnitude and sign of long-range correlated time series: Decomposition and surrogate signal generation. *Physical Review E*, 93(4), 042201, 2016. (Gómez-Extremera et al., 2016)
- M. Gómez-Extremera, P. Carpena and P. A. Bernaola-Galván. Magnitude and sign decomposition of long-range fractal correlated time series. *Proceedings Itise 2015*, 161-170, ISBN: 978-84-16292-20-2.
- P. A. Bernaola-Galván, M. Gómez-Extremera, P. Carpena and A. R. Romance. Correlations in magnitude series to assess nonlinearities: Application to multifractal models and heartbeat fluctuations. *Physical Review E*, 2017. (Bernaola-Galván et al., 2017)
- P. Carpena, M. Gómez-Extremera, C. Carretero-Campos, P. A. Bernaola-Galván and A. V. Coronado. Spurious Results of Fluctuation Analysis Techniques in Magnitude and Sign Correlations. *Entropy*, 19(6), 261, 2017. (Carpena et al., 2017)

D.2. Contribuciones

- Correlations in the magnitude time series as an index of nonlinearity. Application to heartbeat fluctuation during rest and exercise. *First International Summer Institute on Network Physiology*, Como, Italy, 2017.

Bibliografía

- Arto J A. J. Hautala, T. H. Mäkikallio, T. Seppänen, H. V. Huikuri, and M. P. Tulppo. Short-term correlation properties of r-r interval dynamics at different exercise intensity levels. *Clinical physiology and functional imaging*, 23(4):215–223, 2003.
- P. Allegrini, M. Barbi, P. Grigolini, and B.J. West. Dynamical model for dna sequences. *Physical Review E*, 52, 5281, 1995.
- O. Anosov, A. Patzak, Y. Kononovich, and P. B. Persson. High-frequency oscillations of the heart rate during ramp load reflect the human anaerobic threshold. *European journal of applied physiology*, 83(4-5):388–394, 2000.
- S. S. Apostolov, F. M. Izrailev, N. M. Makarov, Z. A. Mayzelis, S. S. Melnyk, and O. V. Usatenko. The signum function method for the generation of correlated dichotomic chains. *J. Phys. A: Math. Theor.*, 41, 17501, 2008.
- A. Arneodo and E. Bacry. Ondelettes, multifractal et turbulence de l’adn aux croissances cristallines. *Diderot editeur arts et sciences, Paris*, 1995.
- A. Arneodo, G. Grasseau, and M. Holschneider. Wavelet transform of multifractals. *Physical review letters*, 61(20):2281, 1988.
- A. Arneodo, B. Audit, N. Decoster, J. F. Muzy, and C. Vaillant. Wavelet based multifractal formalism: applications to dna sequences, satellite images of the cloud structure, and stock market data. In *The science of Disasters*, pages 26–102. Springer, 2002.
- Y. Ashkenazy. Magnitude and sign correlations in heartbeat fluctuations. *Physical Review Letter*, 86:1900–1903, 2001.
- Y. Ashkenazy. Magnitude and sign scaling in power-law correlated time series. *Physica A*, 323:19–41, 2003.
- Y. Ashkenazy, D. R. Baker, H. Gildor, and S. Havlin. Nonlinearity and multifractality of climate change in the past 420,000 years. *Geophysical Research Letters*, 30(22), 2003.

- A. E. Aubert, B. Seps, and F. Beckers. Heart rate variability in athletes. *Sports medicine*, 33(12):889–919, 2003.
- G. Badin and D. I. V. Domeisen. Nonlinear stratospheric variability: multifractal detrended fluctuation analysis and singularity spectra. In *Proc. R. Soc. A*, volume 472, page 20150864. The Royal Society, 2016.
- I. Bartos and I.M. János. Nonlinear correlations of daily temperature records over land. *Processes Geophys.*, 13:571–576, 2006.
- J. B. Bassingthwaite, L. S. Liebovitch, and B. J. West. Properties of fractal phenomena in space and time. In *Fractal physiology*, pages 11–44. Springer, 1994.
- J. Beran. *Statistics for long-memory processes*, volume 61. CRC press, 1994.
- P. Bernaola-Galván, P. Carpena, R. Roman-Roldán, and J.L. Oliver. Study of statistical correlations in dna sequences. *Gene* 300, pages 105–115, 2002.
- P. Bernaola-Galván, J.L. Oliver, M. Hackenberg, A.V. Coronado, P.Ch. Ivanov, and P. Carpena. Segmentation of time series with long-range fractal correlations. *Eur. Phys. J.*, 85, 211, 2012.
- Pedro A. Bernaola-Galván, Manuel Gómez-Extremera, A. Ramón Romance, and Pedro Carpena. Correlations in magnitude series to assess nonlinearities: Application to multifractal models and heartbeat fluctuations. *Phys. Rev. E*, 96:032218, Sep 2017. doi: 10.1103/PhysRevE.96.032218.
- C. Bernard. Les phenomenes de la vie. *Bailliere, Paris*, 879:1, 1878.
- L. Bernardi, F. Valle, M. Coco, A. Calciati, and P. Sleight. Physical activity influences heart rate variability and very-low-frequency components in holter electrocardiograms. *Cardiovascular research*, 32(2):234–237, 1996.
- Marián Boguná, Romualdo Pastor-Satorras, and Alessandro Vespignani. Cut-offs and finite size effects in scale-free networks. *The European Physical Journal B-Condensed Matter and Complex Systems*, 38(2):205–209, 2004.
- Christian Borgs and Roman Kotecký. Finite-size effects at asymmetric first-order phase transitions. *Physical review letters*, 68(11):1734, 1992.
- G. R. H. Sandercock D. A. Brodie. The use of heart rate variability measures to assess autonomic control during exercise. *Scandinavian journal of medicine & science in sports*, 16(5):302–313, 2006.

- R. M. Bryce and K. B. Sprague. Revisiting detrended fluctuation analysis. *Scientific Reports* 2, 315, 2012.
- E. A. Byrne, J. L. Fleg, P. V. Vaitkevicius, J. Wright, and S. W. Porges. Role of aerobic capacity and body mass index in the age-associated decline in heart rate variability. *Journal of Applied Physiology*, 81(2):743–750, 1996.
- A John Camm, Marek Malik, JT Bigger, Günter Breithardt, Sergio Cerutti, Richard J Cohen, Philippe Coumel, Ernest L Fallen, Harold L Kennedy, Robert E Kleiger, et al. Heart rate variability: standards of measurement, physiological interpretation and clinical use. task force of the european society of cardiology and the north american society of pacing and electrophysiology. *Circulation*, 93(5):1043–1065, 1996.
- P. Carpena, J. Aguiar, P. Bernaola-Galván, and C. Carnero-Ruiz. Problems associated with the treatment of conductivity-concentration data in surfactant solutions. *Langmuir* 18, pages 6054–6058, 2002.
- P Carpena, P Bernaola-Galván, M Hackenberg, AV Coronado, and JL Oliver. Level statistics of words: Finding keywords in literary texts and symbolic sequences. *Physical Review E*, 79(3):035102, 2009.
- P. Carpena, M. Gómez-Extremera, C. Carretero-Campos, P. Bernaola-Galván, and A. V. Coronado. Spurious results on fluctuation analysis techniques in magnitude and sign correlations. *Entropy*, 19(6), 2017.
- C. Carretero-Campos, P. Bernaola-Galván, P. Ch. Ivanov, and P. Carpena. Phase transitions in the first-passage time of scale-invariant correlated processes. *Physical Review E*, 85, 011139, 2012.
- Z. Chen. Effect of nonstationarities on detrended fluctuation analysis. *Physical Review E*, 65, 041107, 2002.
- Z. Chen. Effect of nonlinear filters on detrended fluctuation analysis. *Physical Review E*, 71, 011104, 2005.
- R. R. Coifman, M. V. Wickerhauser Y. Meyer, et al. Size properties of wavelet packets. *Wavelets and their applications*, pages 453–470, 1992.
- Christopher R Cole, Eugene H Blackstone, Fredric J Pashkow, Claire E Snader, and Michael S Lauer. Heart-rate recovery immediately after exercise as a predictor of mortality. *New England Journal of Medicine*, 341(18):1351–1357, 1999.

- Ana V Coronado and Pedro Carpena. Size effects on correlation measures. *Journal of biological physics*, 31(1):121–133, 2005.
- Álvaro Corral. Dependence of earthquake recurrence times and independence of magnitudes on seismicity history. *Tectonophysics*, 424(3):177–193, 2006.
- K. P. Davy, W. L. Willis, and D. R. Seals. Influence of exercise training on heart rate variability in post-menopausal women with elevated arterial blood pressure. *Clinical Physiology and Functional Imaging*, 17(1):31–40, 1997.
- F. A. B. F. de Moura and M. L. Lyra. Delocalization in the 1d anderson model with long-range correlated disorder. *Phys. Rev. Lett.*, 81:3735–3738, 1998.
- Agamemnon Despopoulos and Stefan Silbernagl. *Color atlas of physiology*. Thieme., 2003.
- J. G. Dong. The role of heart rate variability in sports physiology (review). *Experimental and therapeutic medicine*, 11(5):1531–1536, 2016.
- G. Erlebacher, M. Y. Hussaini, and L. M. Jameson. *Wavelets: theory and applications*. Oxford University Press on Demand, 1996.
- Gernot Ernst. *Heart rate variability*. Springer, 2014.
- U. Frisch and G. Parisi. On the singularity structure of fully developed turbulence. 1985.
- A. L. Goldberger, L. A. N. Amaral and J. M. Hausdorff, P. Ch. Ivanov, C. K. Peng, and H. E. Stanley. Fractal dynamics in physiology: alterations with disease and aging. *Proceedings of the National Academy of Sciences*, 99(suppl 1):2466–2472, 2002.
- Jeffrey J Goldberger. Sympathovagal balance: how should we measure it? *American Journal of Physiology-Heart and Circulatory Physiology*, 276(4):H1273–H1280, 1999.
- R. L. Goldsmith, J. T. Bigger, R. C. Steinman, and J. L. Fleiss. Comparison of 24-hour parasympathetic activity in endurance-trained and untrained young men. *Journal of the American College of Cardiology*, 20(3):552–558, 1992.
- Manuel Gómez-Extremuera, Pedro Carpena, Plamen Ch Ivanov, and Pedro A Bernaola-Galván. Magnitude and sign of long-range correlated time series: Decomposition and surrogate signal generation. *Physical Review E*, 93(4):042201, 2016.
- P. Goupillaud, A. Grossmann, and J. Morlet. Cycle-octave and related transforms in seismic signal analysis. *Geoexploration*, 23(1):85 – 102, 1984.

- I. S. Gradshteyn and I. M. Ryzhik. *Table of integrals, series, and products*. Academic press, 2014.
- A. Grossmann and J. Morlet. Mathematics and physics, lectures on recent results. *Streit World Scientific, Singapour*, 1985.
- Michael Hackenberg, Pedro Carpena, Pedro Bernaola-Galván, Guillermo Barturen, Ángel M Alganza, and José L Oliver. Wordcluster: detecting clusters of dna words and genomic elements. *Algorithms for Molecular Biology*, 6(1):2, 2011.
- Michael Hackenberg, Antonio Rueda, Pedro Carpena, Pedro Bernaola-Galván, Guillermo Barturen, and José L Oliver. Clustering of dna words and biological function: a proof of principle. *Journal of theoretical biology*, 297:127–136, 2012.
- K. S. Heffernan, C. A. Fahs, K. K. Shinsako, S. Y. Jae, and B. Fernhall. Heart rate recovery and heart rate complexity following resistance exercise training and detraining in young men. *American Journal of Physiology-Heart and Circulatory Physiology*, 293(5):H3180–H3186, 2007.
- R. Herrmann. *Fractional calculus: an introduction for physicists*. World Scientific, 2014.
- J. R. M. Hosking. Fractional differencing. *Biometrika*, 68:165–176, 1981.
- K. Hu. Effect of trends on detrended fluctuation analysis. *Physical Review E*, 64, 011114, 2001.
- H. E. Hurst. Long-term storage capacity of reservoirs. *Trans. Amer. Soc. Civil Eng.*, 116: 770–808, 1951.
- B. W. Hyndman. The role of rhythms in homeostasis. *Biological Cybernetics*, 15(4): 227–236, 1974.
- Katsuji Imai, Hideyuki Sato, Masatsugu Hori, Hideo Kusuoka, Hitoshi Ozaki, Hiroshi Yokoyama, Hiroshi Takeda, Michitoshi Inoue, and Takenobu Kamada. Vagally mediated heart rate recovery after exercise is accelerated in athletes but blunted in patients with chronic heart failure. *Journal of the American College of Cardiology*, 24(6):1529–1535, 1994.
- P. Ch. Ivanov. Multifractality in human heartbeat dynamics. *Nature*, 399:461–465, 1999a.
- P. Ch. Ivanov, M. G. Rosenblum, C. K. Peng, J. Mietus, et al. Scaling behaviour of heartbeat intervals obtained by wavelet-based time-series analysis. *Nature*, 383(6598): 323, 1996.

- P Ch Ivanov, Armin Bunde, LA Nunes Amaral, Shlomo Havlin, Janice Fritsch-Yelle, Roman M Baevsky, H Eugene Stanley, and Ary L Goldberger. Sleep-wake differences in scaling behavior of the human heartbeat: Analysis of terrestrial and long-term space flight data. *EPL (Europhysics Letters)*, 48(5):594, 1999.
- P.Ch. Ivanov. Analysis of terrestrial and long-term spacial flight data. *Europhysics Lett*, 48(5):594–600, 1999b.
- Nikhil Iyengar, CK Peng, Raymond Morin, Ary L Goldberger, and Lewis A Lipsitz. Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 271(4):R1078–R1084, 1996.
- F.M. Izrailev, A.A. Krokhin, N.M. Makarov, and O.V. Usatenko. Generation of correlated binary sequences from white noise. *Physical Review E*, 76, 027701, 2007.
- M Javorka, I Zila, T Balharek, and K Javorka. Heart rate recovery after exercise: relations to heart rate variability and complexity. *Brazilian Journal of Medical and Biological Research*, 35(8):991–1000, 2002.
- R. A. Johnson, D. W. Wichern, et al. *Applied multivariate statistical analysis*, volume 5. Prentice hall Upper Saddle River, NJ, 2002.
- T. Kalisky, Y. Ashkenazy, and S. Havlin. Volatility of linear and nonlinear time series. *Physical Review E*, 72, 011913, 2005.
- J. A. Kanaley, S. Goulopoulou, R. M. Franklin, T. Baynard, M. E. Holmstrup, J. R. Carhart, R. S. Weinstock, and B. Fernhall. Plasticity of heart rate signaling and complexity with exercise training in obese individuals with and without type 2 diabetes. *International journal of obesity (2005)*, 33(10):1198, 2009.
- J.W Kantelhardt. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A*, 316, 87, 2002.
- J.W. Kantelhardt, S. Russ, A. Bunde, S. Havlin, and I. Webman. Comment on "delocalization in the 1d anderson model with long-range correlated disorder". *Phys. Rev. Lett.*, 84, 198, 2000.
- R. Karasik, N. Sapir, Y. Ashkenazy, P. Ch. Ivanov, I. Dvir, P. Lavie, and S. Havlin. Correlation differences in heartbeat fluctuations during rest and exercise. *Physical Review E*, 66(6):062902, 2002.

- S. Hod & U. Keshet. Phase transition in random walks with long-range correlations. *Physical Review E*, 70, 015104(R), 2004.
- Peter Krstrup, Magni Mohr, Tommas Amstrup, Torben Rysgaard, Johnny Johansen, Adam Steensberg, Preben K Pedersen, and Jens Bangsbo. The yo-yo intermittent recovery test: physiological response, reliability, and validity. *Medicine & Science in Sports & Exercise*, 35(4):697–705, 2003.
- Timo T Laitio, Heikki V Huikuri, Erkki SH Kentala, Timo H Mäkikallio, Jouko R Jalonen, Hans Helenius, Kaisa Sariola-Heinonen, Sinikka Yli-Mäyry, and Harry Scheinin. Correlation properties and complexity of perioperative rr-interval dynamics in coronary artery bypass surgery patients. *Anesthesiology: The Journal of the American Society of Anesthesiologists*, 93(1):69–80, 2000.
- W. C. Levy, M. D. Cerqueira, G. D. Harp, K. A. Johannessen, I. B. Abrass, R. S. Schwartz, and J. R. Stratton. Effect of endurance exercise training on heart rate variability at rest in healthy young and older men. *The American journal of cardiology*, 82(10):1236–1241, 1998.
- M. J. Lewis and A. L. Short. Exercise and cardiac regulation: what can electrocardiographic time series tell us? *Scandinavian journal of medicine & science in sports*, 20(6):794–804, 2010.
- Q. Li, Z. Fu, N. Yuan, and F. Xie. Effects of non-stationarity on the magnitude and sign scaling in the multi-scale vertical velocity increment. *Physica A*, 410:9–16, 2014.
- Wentian Li and Kunihiko Kaneko. Long-range correlation and partial $1/f\alpha$ spectrum in a noncoding dna sequence. *EPL (Europhysics Letters)*, 17(7):655, 1992.
- Michael J Lipinski, George W Vetrovec, and Victor F Froelicher. Importance of the first two minutes of heart rate recovery after exercise treadmill testing in predicting mortality and the presence of coronary artery disease in men. *The American journal of cardiology*, 93(4):445–449, 2004.
- Y.H. Liu. Statistical properties of the volatility of price fluctuations. *Physical Review E*, 60:1390–1400, 1999.
- S. B. Lowen and M. C. Teich. *Fractal-based point processes*, volume 366. John Wiley & Sons, 2005.
- Q.D.Y. Ma. Effects of extreme data loss on detrended fluctuation analysis. *Physical Review E*, 81, 031101, 2010.

- Timo H Mäkikallio, Tiina Ristimäe, KE Juhani Airaksinen, Chung-Kang Peng, Ary L Goldberger, and Heikki V Huikuri. Heart rate dynamics in patients with stable angina pectoris and utility of fractal and complexity measures. *The American journal of cardiology*, 81(1):27–31, 1998.
- H. A. Makse, S. Havlin, M. Schwartz, and H. E. Stanley. Method for generating long-range correlations for large systems. *Physical Review E*, 53, 5445, 1996a.
- H.A. Makse, G.W. Davies, P.Ch. Ivanov S. Havlin, P.R. King, S. Prakash, and H.E. Stanley. Long-range correlation in permeability fluctuations in porous rock. *Physical Review E*, 54:3129–3134, 1996b.
- H.A. Makse, G.W. Davies, P.Ch. Ivanov S. Havlin, P.R. King, S. Prakash, and H.E. Stanley. Pattern formation in sedimentary rock. *Physica A*, 233:587–605, 1996c.
- M. Malik and A. J. Camm. *Heart rate variability*. Futura Publishing Company Armonk, NY, 1995.
- S. Mallat, W. L. Hwang, and W. Liang. Singularity detection and processing with wavelets. *IEEE transactions on information theory*, 38(2):617–643, 1992.
- B. B. Mandelbrot. Stochastic models for the earth's relief, the shape and the fractal dimension of the coastlines, and the number-area rule for islands. *Proceedings of the National Academy of Sciences*, 72(10):3825–3828, 1975.
- B. B. Mandelbrot. Fractals: form, chance and dimension. *Fractals: form, chance and dimension. San Francisco (CA, USA): WH Freeman & Co., 16+ 365 p., 1, 1979.*
- B. B. Mandelbrot and R. Pignoni. *The fractal geometry of nature*, volume 1. WH freeman New York, 1983.
- B. B. Mandelbrot and J. W. van Ness. Fractional brownian motions, fractional noises and applications. *SIAM Rev*, 10, No.4:422–437, 1968.
- B. B. Mandelbrot and J. R. Wallis. Noah, joseph and operational hydrology. *Water Resources Res*, 4, No 5:909–918, 1968a.
- B. B. Mandelbrot and J. R. Wallis. Robustness of the rescaled range r/s and the measurement of non-cyclic long-run statistical dependence. *Water Resources Res*, 5:967–988, 1968b.

- B. B. Mandelbrot and J. R. Wallis. Computer experiments with fractional gaussian noises. *Water Resources Res*, 5, No. 1:228–267, 1969a.
- B. B. Mandelbrot and J. R. Wallis. Some long-run properties of geophysical records. *Water Resources Res*, 5:321–340, 1969b.
- J. M. Martín-González and J. M. García-Manso. Wavelet transforms in sport: Application to biological time series. In *Wavelet Transforms and Their Recent Applications in Biology and Geoscience*. InTech, 2012.
- Alexander J McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative risk management: Concepts, techniques and tools*. Princeton university press, 2015.
- E. W. Ng and M. Geller. A table of integrals of the error functions. *Journal of Research of the National Bureau of Standards B*, 73(1):1–20, 1969.
- Miguel Ortuno, Pedro Carpena, Pedro Bernaola-Galván, Enrique Munoz, and Andrés M Somoza. Keyword detection in natural languages and dna. *EPL (Europhysics Letters)*, 57(5):759, 2002.
- P.Ch. Ivanov P. Carpena, P. Bernaola-Galván and H.E. Stanley. Metal-insulator transition in chains with correlated disorder. *Nature*, 418, 955, 2002.
- C. K. Peng, S. V. Buldyrev, A. L. Goldberger, S. Havlin, F. Sciortino, M. Simons, and H. E. Stanley. Long-range correlations in nucleotides. *Nature*, 356:168–170, 1992.
- C-K Peng, SV Buldyrev, AL Goldberger, S Havlin, M Simons, and HE Stanley. Finite-size effects on long-range correlations: Implications for analyzing dna sequences. *Physical Review E*, 47(5):3730, 1993a.
- C. K. Peng, J. Mietus, J. M. Hausdorff, S. Havlin, H. E. Stanley, and A. L. Goldberger. Long-range anticorrelations and non-gaussian behavior of the heartbeat. *Physical review letters*, 70(9):1343, 1993b.
- C. K. Peng, S. V. Buldyrev, J. M. Hausdorff, S. Havlin, J. E. Mietus, M. Simons, H. E. Stanley, and A. L. Goldberger. Non-equilibrium dynamics as an indispensable characteristic of a healthy biological system. *Integrative Physiological and Behavioral Science*, 29(3):283–293, 1994a.
- C. K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger. Mosaic organization of dna nucleotides. *Physical review e*, 49(2):1685, 1994b.

- C-K Peng, Shlomo Havlin, H Eugene Stanley, and Ary L Goldberger. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 5(1):82–87, 1995.
- P. Perakakis, M. Taylor, E. Martinez-Nieto, I. Revithi, and J. Vila. Breathing frequency bias in fractal analysis of heart rate variability. *Biological psychology*, 82(1):82–88, 2009.
- Edgar E Peters. *Fractal market analysis: applying chaos theory to investment and economics*, volume 24. John Wiley & Sons, 1994.
- Sirkku M Pikkujäämsä, Timo H Mäkikallio, Leif B Sourander, Ismo J Räihä, Pauli Puukka, Jarmo Skyttä, Chung-Kang Peng, Ary L Goldberger, and Heikki V Huikuri. Cardiac interbeat interval dynamics from childhood to senescence. *Circulation*, 100(4):393–399, 1999.
- M. M. Platisa, S. Mazic, Z. Nestorovic, and V. Gal. Complexity of heartbeat interval series in young healthy trained and untrained men. *Physiological measurement*, 29(4):439, 2008.
- G. Rangarajan and M.Z. Ding. Integrated approach to the assessment of long range correlation in time series data. *Physical Review E*, 61, 4991, 2000.
- S. Sarmiento, J. M. García-Manso, J. M. Martín-González, D. Vaamonde, J. Calderón, and M. E. Da Silva-Grigoletto. Heart rate variability during high-intensity exercise. *Journal of Systems Science and Complexity*, 26(1):104–116, 2013.
- WILLIAM M Savin, DENNIS M Davidson, and WILLIAM L Haskell. Autonomic contribution to heart rate recovery from exercise in humans. *Journal of Applied Physiology*, 53(6):1572–1575, 1982.
- T. Schreiber and A. Schmitz. Surrogate time series. *Physica D: Nonlinear Phenomena*, 142(3):346–382, 2000.
- A. Y. Schumann. *Fluctuations and synchronization in complex physiological systems*. PhD thesis, Halle (Saale), Martin-Luther-Universität Halle-Wittenberg, Diss., 2010, 2010.
- D. R. Seals and P. B. Chase. Influence of physical training on heart rate variability and baroreflex circulatory control. *Journal of Applied Physiology*, 66(4):1886–1895, 1989.
- M. P. Tulppo, R.L. Hughson, T. H. Mäkikallio, K. E. J. Airaksinen, T. Seppänen, and H. V. Huikuri. Effects of exercise and passive head-up tilt on fractal and complexity pro-

- perties of heart rate dynamics. *American Journal of Physiology-Heart and Circulatory Physiology*, 280(3):H1081–H1087, 2001.
- M. P. Tulppo, A. J. Hautala, T. H. Mäkikallio, R. T. Laukkanen, S. Nissilä, R. L. Hughson, and H. V. Huikuri. Effects of aerobic training on heart rate dynamics in sedentary subjects. *Journal of Applied Physiology*, 95(1):364–372, 2003.
- O.V. Usatenko, S.S. Melnik, S.S. Apostolov, N.M. Makarov, and A.A. Krokhin. Iterative method for generating correlated binary sequences. *Physical Review E*, 90, 053305, 2014.
- B Van Der Pol and J Van der Mark. Lxxii. the heartbeat considered as a relaxation oscillation, and an electrical model of the heart. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 6(38):763–775, 1928.
- R. F. Voss and J. Clarke. "1/f noise" in music: Music from 1/f noise. *The Journal of the Acoustical Society of America*, 63(1):258–263, 1978.
- Richard F Voss. Evolution of long-range fractal correlations and 1/f noise in dna base sequences. *Physical review letters*, 68(25):3805, 1992.
- M. Weippert, M. Kumar, S. Kreuzfeld, D. Arndt, A. Rieger, and R. Stoll. Comparison of three mobile devices for measuring r–r intervals and heart rate variability: Polar s810i, suunto t6 and an ambulatory ecg system. *European journal of applied physiology*, 109(4):779–786, 2010.
- W. Wyss. Fractional noise. *Foundations of Physics Letters*, 4(3):235–246, 1991.
- L. Xu. Quantifying signals with powerlaw correlations: A comparative study of detrended fluctuation analysis and detrended moving average techniques. *Physical Review E*, 71, 051101, 2005.
- Y. Xu, Q. D. Y. Ma, D. T. Schmitt, P. Bernaola-Galván, and P. C. Ivanov. Effects of coarse-graining on the scaling behavior of long-range correlated and anti-correlated signals. *Physica A*, 390:23–24, 2010.
- L. Zhu. Magnitude and sign correlations in conductance fluctuations of horizontal oil water two-phase flow. *J. Phys. Conf. Ser*, 364, 012067, 2012.